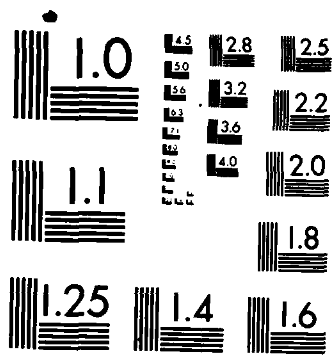
AD-A161 355 EXPLORATION OF THE MAXIMUM ENTROPY/OPTIMAL PROJECTION APPROACH TO CONTROL (U) HARRIS CORP MELBOURNE FL GOVERNMENT AEROSPACE SYSTEMS DIV D C HYLAND ET AL 1/3

UNCLASSIFIED FEB 85 AFOSR-TR-85-0966 F49620-84-C-0015 F/G 13/13 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

EXPLORATION OF THE MAXIMUM ENTROPY/OPTIMAL PROJECTION APPROACH TO CONTROL DESIGN SYNTHESIS FOR LARGE SPACE STRUCTURES

D. C. Hyland
D. S. Bernstein

Harris Corporation
Government Aerospace Systems Division
Melbourne, Florida 32902

Annual Report No. 1
Contract No. F49620-84-C-0015

Prepared for

Air Force Office of Scientific Research
Directorate of Aerospace Sciences
Bolling AFB, D.C. 20332

February 1985

HARRIS

HARRIS CORPORATION GOVERNMENT AEROSPACE SYSTEMS DIVISION
P.O. BOX 94000, MELBOURNE, FLORIDA 32902, (305) 727-5115

2

EXPLORATION OF THE MAXIMUM ENTROPY/OPTIMAL

PROJECTION APPROACH TO CONTROL DESIGN

SYNTHESIS FOR LARGE SPACE STRUCTURES

D. C. Hyland
D. S. Bernstein

Harris Corporation
Government Aerospace Systems Division
Melbourne, Florida 32902

Annual Report No. 1
Contract No. F49620-84-C-0015

Prepared for

Air Force Office of Scientific Research
Directorate of Aerospace Sciences
Bolling AFB, D.C. 20332

February 1985

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited. Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 8 5 - 0 9 3 6
6a. NAME OF PERFORMING ORGANIZATION Harris Corporation Government Aerospace Div. — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NA
6c. ADDRESS (City, State and ZIP Code) MS 102/4124 Melbourne, Fl 32901	7b. ADDRESS (City, State and ZIP Code) BOLLING AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Office of Scientific Research — 8b. OFFICE SYMBOL (If applicable) AFOSR/NA	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-84-C-0015
8c. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB Washington, DC 20322	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102 F PROJECT NO.: 2307 TASK NO.: B1 WORK UNIT NO.:
11. TITLE (Include Security Classification) Max.Entropy/Optimal Projection For LSS Control	
12. PERSONAL AUTHOR(S) Hyland, David, Charles Bernstein, Dennis, S.	
13a. TYPE OF REPORT Annual	13b. TIME COVERED FROM 02/01/84 TO 02/28/85
14. DATE OF REPORT (Yr., Mo., Day) 1985, February	15. PAGE COUNT 240
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Stochastic processes; parameter uncertainty; Maximum Entropy Formalism; Quadratic Optimization; Reduced-Order Dynamic Compensation,
22	02		
12	01		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Increased interest in deploying large flexible spacecraft has focussed attention on active structural control techniques to achieve crucial advances in vibration suppression, pointing accuracy and shape control. The extreme complexity of such systems, and the lack of accurate finite-element structural models present severe control-design challenges which were extensively documented by previous Government research Programs. Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis is a rigorous new approach to this class of problems. Inspired by Statistical Energy Analysis, a branch of dynamic modal analysis developed for analyzing acoustic vibration problems, its present stage of development embodies a fundamental generalization of classical steady-state Kalman filter and linear-quadratic-Gaussian ~~(LQG)~~ optimal control theory. Keywords:

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL ANTHONY K AMOS	22b. TELEPHONE NUMBER (Include Area Code) (202) 767-4937 — 22c. OFFICE SYMBOL AFOSR/NA

Although LQG theory is an effective tool for optimally quantifying performance/sensor-resolution and performance/actuation-level tradeoffs, it suffers from two fundamental defects which severely limit its usefulness in practice. These defects are addressed by the optimal projection/maximum entropy design methodology:

1. Whereas the dimension of an LQG controller must equal that of the controlled plant, optimal projection design characterizes the quadratically optimal controller of fixed dimension <u>less</u> than that of the plant in accordance with implementation constraints (e.g., reliability, complexity or real-tine computing capability).

2. Whereas LQG presumes exact knowledge of each and every parameter appearing in the state-space plant description, maximum entropy modelling provides a stochastic plant model which admits ignorance with regard to parameter values in accordance with unavoidable plant modelling errors.

The principal aim of this research is to investigate the impact of flexible spacecraft structural modelling uncertainties and dimensionality on active control design through the exploitation of the maximum entropy stochastic modelling approach together with the optimal projection formulation of quadratically optimal, fixed-order dynamic compensation. Within this context, the main goal is to implement and demonstrate a unified design synthesis technique which permits direct design in the face of incomplete system information while preserving a well-defined notion of optimality.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

APPENDICES

SECTION 1.0

INTRODUCTION

1.0 INTRODUCTION

Increased interest in deploying large flexible spacecraft has focussed attention on active structural control techniques to achieve crucial advances in vibration suppression, pointing accuracy and shape control. The extreme complexity of such systems and the lack of accurate finite-element structural models present severe control-design challenges which were extensively documented by DARPA's ACOSS Program. Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis is a rigorous new approach to this class of problems conceived by Dr. D. C. Hyland in [2-16]. Inspired by Statistical Energy Analysis ([1]), a branch of dynamic modal analysis developed for analyzing acoustic vibration problems, its present stage of development ([2-30], see Appendices A-F), embodies a mathematically rigorous, fundamental generalization of classical steady-state Kalman filter and linear-quadratic-Gaussian (LQG) optimal control theory.

Although LQG theory is an effective tool for optimally quantifying performance/sensor-resolution and performance/actuation-level tradeoffs, it suffers from two fundamental defects which severely limit its usefulness in practice. These defects are remedied by the optimal projection/maximum entropy design methodology:

1. Whereas the dimension of an LQG controller must equal that of the controlled plant, optimal projection design characterizes the quadratically optimal controller of fixed dimension <u>less</u> than that of the plant in accordance with implementation constraints (e.g., reliability, complexity or real-time computing capability).

2. Whereas LQG presumes exact knowledge of each and every parameter appearing in the state-space plant description, maximum entropy modelling provides a stochastic plant model which admits ignorance with regard to parameter values in accordance with unavoidable plant modelling errors.

With regard to the latter item, it should be stressed that one of the major problems in designing high-performance control systems is that of robustness, i.e., the ability of the controller to tolerate errors in the plant model upon which its design is predicated. Maximum entropy modelling directly addresses this problem by incorporating into the dynamic model a representation of ignorance (i.e., uncertainty) regarding physical parameters. Roughly speaking, the idea behind the approach is to use a probabilistic representation of each imperfectly known plant parameter. The quadratically optimal control system designed in the presence of this probabilistic model is automatically desensitized to actual parameter variations when the control system is implemented. The overall control-design procedure thus avoids laborious trial-and-error post-design "tweaking."

The validity of the above claims has been demonstrated on a series of representative structural models of increasing complexity. Beginning with an illustrative simply supported beam for conceptual validation ([5]), the theory has since been applied to both a 20-state version of the Draper Model #2 ([19], Appendix A) and a 16-state model of NASA's SCOLE experiment.

Although the early development of the optimal projection/maximum entropy approach was documented in numerous conference proceedings, a series of manuscripts has recently entered the realm of archival publications. Reference [23] (Appendix B) recently appeared in the November issue of the IEEE Transactions on Automatic Control and reference [29] (Appendix C) has been accepted as a full paper by this journal. Reference [28] (Appendix D), which rigorously extends the results of [23] to distributed parameter systems, is scheduled to appear in SIAM Journal on Control and Optimization. In addition, a series of manuscripts ([31]-[34], see also Appendix E) is under preparation detailing the results of the initial 12-month contract period. These papers, which will be submitted to both conferences and journal publications, will considerably broaden the theoretical scope of the OP/ME approach and further demonstrate its practical applicability.

1.1 Study Overview

The principal aim of this effort is to investigate the impact of flexible spacecraft structural modelling uncertainties and dimensionality on active control design through the exploitation of the minimal data/maximum entropy

stochastic modelling approach together with the optimal projection formulation of quadratically optimal, fixed-order dynamic compensation. Within this context, the principal aim is to implement and demonstrate a unified design synthesis technique which permits direct design in the face of incomplete system information while preserving a well-defined notion of optimality. The tasks required to accomplish the goals of this contract are discussed in detail within the original technical proposal and are briefly summarized as follows:

Task 1: Implement fully mechanized construction of the uncertainty operators employed by the maximum entropy modelling approach. These uncertainty operators would be based on very general descriptions of structural parameter uncertainties as represented in a variety of vector bases. The purpose of this task is to facilitate interaction between control-system designers and structural analysts by streamlining the stochastic modelling process and extending maximum entropy modelling theory and its mechanizations to the treatment of uncertainty in general classes of physical structural parameters.

Task 2: Develop and test techniques for numerically solving the fixed-order dynamic-compensation optimality conditions given in [7]. These computational techniques would be implemented in a "user-friendly" form to facilitate the design process and would be capable of handling design models of high order (10^2 states). This task is planned in two stages:

2.A. Develop and test techniques for solving the optimal projection equations for deterministically parametered system models and/or maximum entropy stochastic models in the coherent regime. Such techniques, by employing suitable variants of standard LQG software modules, should, therefore, be capable of handling problems of modest dimension (≤ 20 modes).

2.B. Develop advanced, relaxation-type solution techniques that exploit the crucial incoherence and isotropy effects of maximum entropy models. These techniques would incorporate the developments in 2.A while permitting treatment of very large dimensional systems.

Task 3: Demonstrate the design capabilities developed in Tasks 1 and 2 on the EPAR and Draper Model #2 spacecraft models as well as on simple high-order examples (beams, plates, and the like). As argued in [8], the mechanizations to be developed under Task 2 above would constitute basic computational tools possessing a variety of unique features and could play a pivotal role in both structural and control-system design.

Our goal within the first year of this study* was to complete Task 2.A and initiate Task 3. This and more have been accomplished. In particular, Task 1 was addressed, and new theoretical and practical results were obtained for the basic maximum entropy approach. Task 2.A is largely complete, and quadratically optimal, low-order control designs were demonstrated for a version of the CSDL Model #2. In addition, the combined optimal projection/maximum entropy design method has been developed, implemented, and tested on a spacecraft example problem of intense current interest.

The above developments are outlined in the following sections of this report. Since many of these investigations have appeared or will shortly appear in the technical literature, we confine the presentation to a discussion of results with explanatory material as needed.

Section 2.0 reviews recent theoretical advances in maximum entropy modelling. Work prior to the present effort considered a somewhat restricted class of parameter uncertainties occurring in structural modelling and induced a maximum entropy stochastic model from a set of parameter statistical data consisting of the relaxation time-scales. Although the relaxation times constitute the minimum data set needed to induce a well-defined probability model via the maximum entropy principle, their use as basic descriptors of the degree of parameter uncertainty differs somewhat from traditional engineering practice. To close this gap, Section 2.0 presents the maximum entropy stochastic model for which the available parameter data is given in terms of bounds on parameter deviations from nominal values. This new formulation readily extends to a very general class of parameter uncertainties and thus considerably generalizes earlier results.

---

*See "Addendum to Section 5.0 of the Contract Proposal," 14 October, 1983, which gives the revised work schedule.

In addition, we consider what is needed in the formulation in order to provide an a priori guarantee of performance over the stipulated range of parameter variations. It turns out that such a guarantee is afforded by a relatively innocuous and straightforward augmentation of the basic maximum entropy model. It should also be pointed out that the extended maximum entropy formulation can be applied to nonlinearities obeying sector inequalities and to structural perturbations which are important to fault tolerance/reliability questions.

Section 3.0 deals with the optimal projection approach to designing quadratically optimal, reduced-order dynamic controllers for high-order systems. First, we present the fundamental optimal projection equations for finite-dimensional systems with deterministic parameters and discuss various computational algorithms for their solution. Optimal projection design results for a variant of the CSDL Model #2 spacecraft control problem are reviewed and compared with the results of various suboptimal controllers obtained by means of order-reduction schemes. These results demonstrate the completion of Task 2.A.

Recent theoretical extensions of the original optimal projection formulation are also pointed out in this section. First, its extension to distributed parameter systems has been carried on in [28] and, second, the optimal projection equations for finite-dimensional systems characterized by a general maximum entropy stochastic model reflecting uncertainties in all system matrices are given in [32]. The resulting design equations provide a unified basis for subsequent practical implementations.

In addition, the essential optimal projection idea gives rise to new and more powerful approaches to two problems of longstanding importance in systems theory: optimal model reduction and quadratically optimal reduced-order state estimation. The fundamental optimal projection equations for these two problems are given in [29] and [30] in Appendices C and F.

Section 4.0 returns to the main theme by outlining the combined optimal projection/maximum entropy (OP/ME) design approach. The results summarized in Sections 2.0 and 3.0 are applied to a structural system having uncertainties in the stiffness matrix. Computational techniques are presented and illustrated by very recent results obtained for the NASA/IEEE Design Challenge configuration.

SECTION 2.0

MAXIMUM ENTROPY MODELLING

2.0 MAXIMUM ENTROPY MODELLING

2.1 Review of the Maximum Entropy Modelling Philosophy

To set the stage for subsequent developments we first review the basic scope and philosophy of our approach to modelling uncertainty and the results obtained for dynamic systems with parameter uncertainties.

First, as to the general scope of these investigations, our considerations will be confined within the context of linear, time-domain, finite-dimensional (but large-order) system models. Moreover, we shall emphasize those design problems, e.g., shape and pointing control of large spacecraft, which require high performance with robustness restricted to the lowest levels consistent with parameter deviations to be expected in practice.

The restriction to linear system design makes sense once one recalls the inevitable tradeoff between robustness and performance. For example, performance requirements during deployment and erection are relatively benign while, at the same time, dynamic nonlinearity due to large angle relative motions of structural components engenders modelling difficulties. This problem calls for a low performance, very robust controller and can be treated effectively by a low performance but inherently dissipative, e.g., positive-real, controller. Thus, the specific impact of parameter uncertainties is not a critical issue in this situation. In contrast, however, the stringent tolerances imposed in many applications upon steady-state figure and line-of-sight (LOS) control demand the highest authority, highest gain control consistent with the minimum permissible robustness levels. Such a control design forces a detailed examination of the influence of specific modelling errors while, due to the performance requirements, elastic deformations are sufficiently small that dynamic nonlinearities may be neglected (at least for steady-state LOS control). Thus, while extension of our approach to nonlinear dynamics would certainly be of interest, the first order of business remains the treatment of modelling uncertainties in linear systems.

To drive the design process we utilize a quadratic functional averaged over the parameter statistical ensemble as our mean-square performance measure. Although various criticisms have been leveled at quadratic optimization, it

entails the simplest and most familiar performance criterion. Although extension of the formulation to more sophisticated costs and performance measures would be of interest, until one is fully capable of dealing with the issues of parameter uncertainty and large dimensionality, this remains a fruitless undertaking. Moreover, the specific consequences of uncertainties even for mean-square optimal design are not fully understood and, therefore, constitute the subject of this work. Most importantly, however, it should be emphasized that quadratic performance criteria (e.g., mean-square pointing error and mean-square reflective surface deformations) are unquestionably meaningful for applications such as large-aperture, spaceborne antenna systems.

The above restrictions to time-domain modelling and quadratic optimization already indicate a fundamental philosophical distinction between this work and the frequency domain L_∞/H_∞ theory currently under development (see [297]-[301]). Moreover, our extensive use of stochastic systems theory stands in stark contrast to the utter determinism of frequency domain theory. In subsequent work we shall evaluate the relative merits of the two approaches as applied to structural control problems. For the present, this matter will not be dwelt upon since we wish to indicate how the maximum entropy approach is itself a fundamental extension of the now classical theory of quadratic optimization for linear systems (e.g., [83]), hereafter referred to as LQG theory.

The evolution of the maximum entropy approach as motivated by problems encountered in the application of LQG theory can be visualized as in Figures 2.1-1 to 2.1-3. Referring to Figure 2.1-1, the model of the plant (assumed to be a structural system here) may be suitably parameterized (e.g., the parameters might be the numerical values of the individual elements of the various system matrices referred to some vector basis) so that both the modelled system and the actual system are defined by their locations within the associated Euclidean space of parameters. LQG-based design approaches implicitly assume that all system maps are known and, consequently, produce a design which is optimal (with respect to a quadratic performance measure) only for a single point (associated with the nominal values) in parameter space. However, due to actual in-mission changes, mathematical modelling errors arising from truncations implicit in the finite element method, etc., all system parameters are not, in fact, known. Thus, to put the matter in the most general terms, a structural model can never encompass the

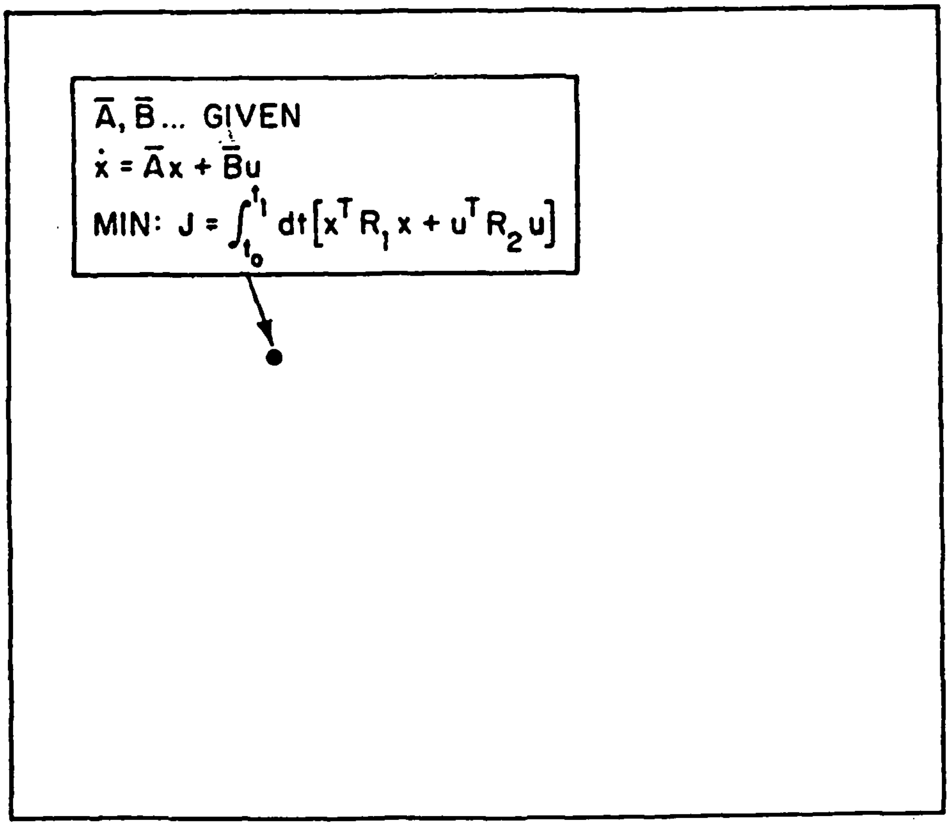


Figure 2.1-1. Modal Parameter Space

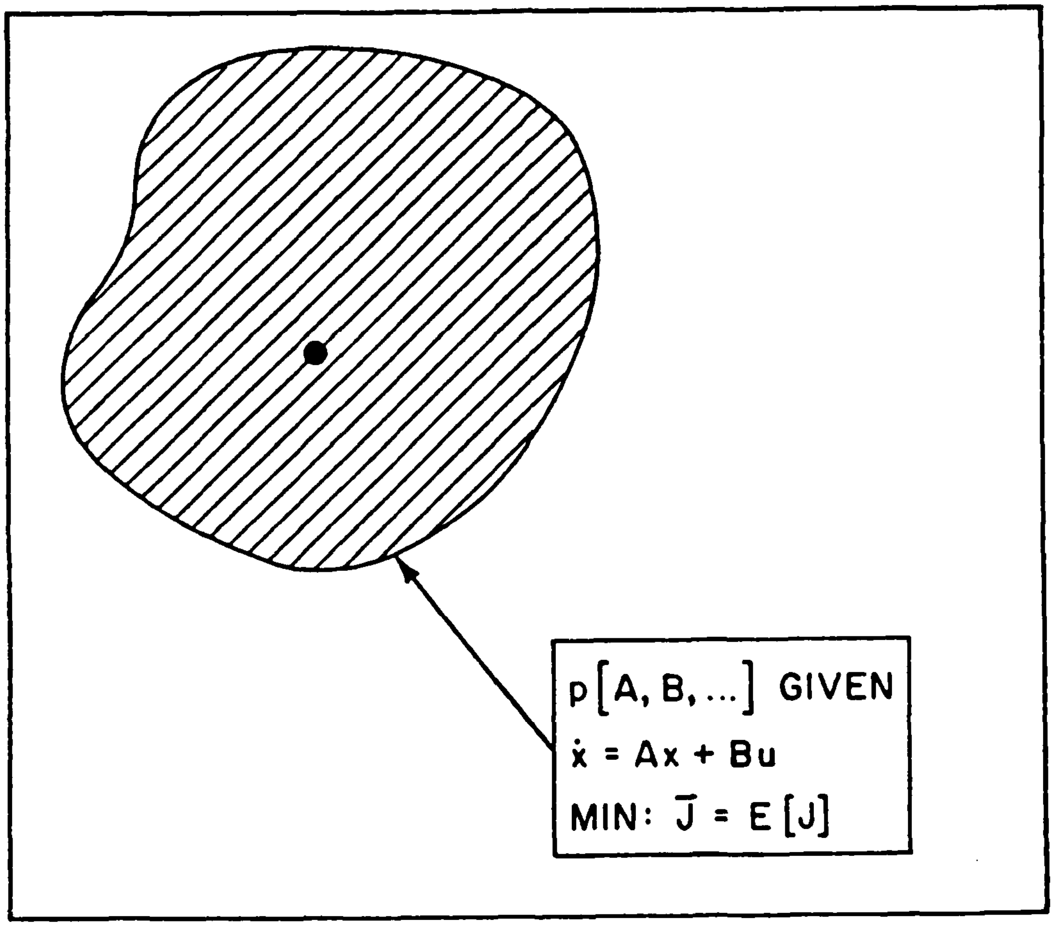


Figure 2.1-2. Probabilistic Description of Parameter Space

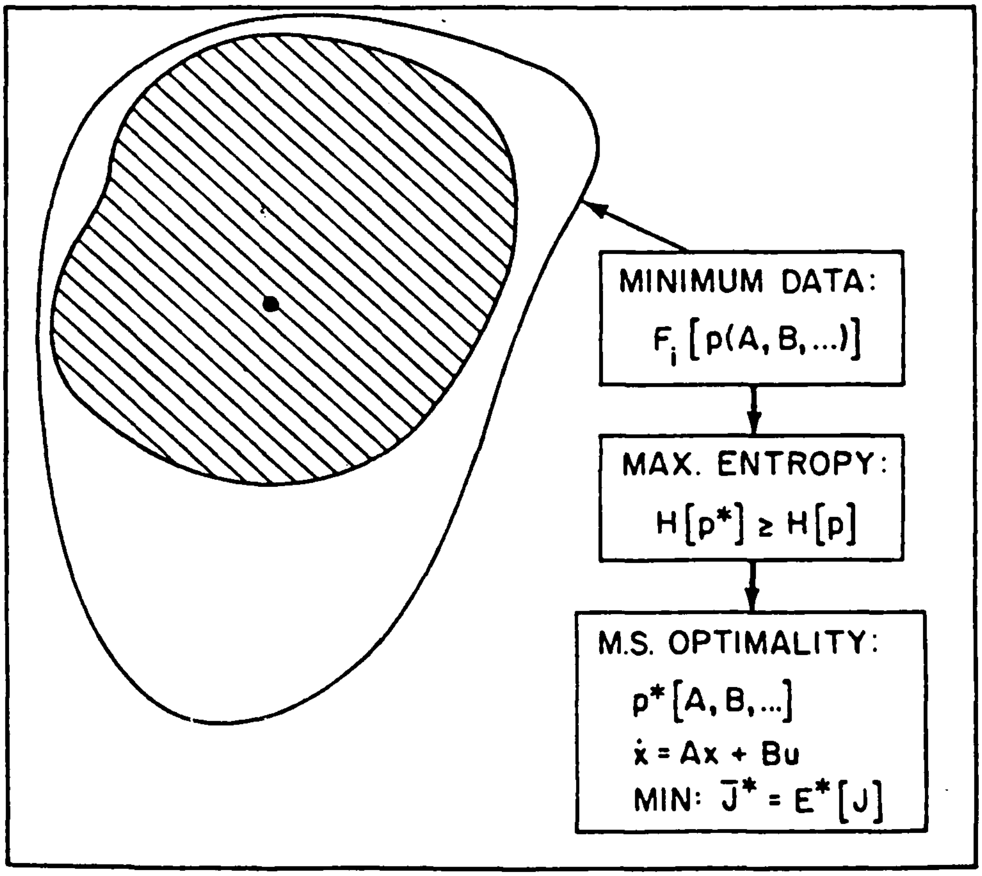


Figure 2.1-3. Maximum Entropy Probability Distribution

"truth"; rather, a model should be regarded as a mathematical statement of what and how much is known. Considered as such, a model must not only specify nominal values but must also contain an admission of prior ignorance regarding parameter deviations from expected values.

An admission of prior uncertainty can indeed be quantified by assuming the parameters to be distributed according to some probability law (see Figure 2.1-2, where the shading indicates the region of significant probability). An important point is that the fundamental concept of probability being employed here is not the traditional relative frequency interpretation but rather the information-theoretic interpretation, i.e., probability as a measure of a priori uncertainty. This matter is elaborated in detail below.

Returning to Figure 2.1-2, one might be tempted to assume that all necessary probability distributions are given and proceed to design a control which is optimal "on the average" by minimizing the expected value of a quadratic cost. Since the quantity of interest is quadratic in the states, the second-moment matrix of the system must be deduced from the given probabilistic description of the multiplicative parameters. This is a longstanding problem in stochastic system theory and the work of Kistner [251] illustrates the significant difficulties of the problem. In general, determination of the second-moment matrix for random multiplicative parameters demands solution of an infinite sequence of systems of ordinary differential equations, i.e., the second-moment equation is not closed.

Such practical difficulties aside, a deeper problem arises: the above standard problem in stochastic systems theory presumes at the outset that a complete probabilistic model is given whereas, in reality, a complete probabilistic description is never available from empirical determinations. Indeed, before undertaking the usual procedures of stochastic systems theory, one must induce a complete probability model from a highly incomplete set of available data. A fundamental logical requirement is that this be done in a manner which avoids inventing data which does not exist! In other words, it is necessary to construct a complete probability assignment which is consistent with the data at hand but admits the greatest possible prior ignorance with regard to all other data. This is the heart of the maximum entropy modelling idea. The appropriate

quantitative procedure has been given by Jaynes ([176-179]): first define a measure of prior ignorance, i.e., the entropy (as in information theory, not thermodynamics), then determine the probability law which maximizes this functional subject to the constraints imposed by available data.

Having overcome the difficulties imposed by incomplete available data in this way, it should be noted that for flexible mechanical systems one can identify a "minimum data set" which is just sufficient to induce any well-defined maximum entropy model. In other words, all admissible sets of available data must include the minimum set, and lack of any element of this minimum set will cause the induced maximum entropy model to "blow up" in certain crucial respects. Since, in practice, one is provided with little or no prior statistical data, it is not only design conservative, but also realistic to acknowledge as available data only the minimum data set.

Thus, as sketched in Figure 2.1-3, our stochastic design approach involves three main stages (see References [3], [8] and [11]). First, the minimum data set is constructed and appropriate numerical values are assigned; next, a maximum-entropy probability model is induced from the minimum data (giving the basic design model); and, finally, a mean-square optimal design is determined under the maximum-entropy statistics. This procedure gives us a mechanism for incorporating incomplete system information within the control design. Moreover, as Figure 2.1-3 suggests, the maximum-entropy model is maximally dispersed in parameter space, and one can guarantee that the resulting design will very greatly reduce the probability of severe performance degradation in the face of parameter deviations.

To recapitulate previous work in maximum entropy modeling and to illustrate how the above procedure is specifically carried out, consider the system

$$\dot{x} = Ax + i\alpha\lambda x + w; \; x \; \underline{C}^n \tag{2.1}$$

where A = nominal value of dynamics matrix ($A + A^*$ assumed stable)

$i\alpha\lambda$ = perturbation in A representing uncertainty in a single parameter

λ = diag $\{\lambda_K\}$, λ_K real

α = real, zero-mean random parameter

w = white noise independent of α having intensity V

Note that for simplicity we consider only a single parameter uncertainty. The key assumption about the perturbation to the nominal dynamics is that it is diagonalizable with purely imaginary eigenvalues, as in earlier work. Also, for simplicity we present the above dynamics in the vector basis diagonalizing the perturbation.

To describe the minimum data set arrived at in [2-11], first note that the equation for the steady-state second-moment matrix

$Q = \underline{E}_{\alpha,w}[xx^*]$ of x is given by

$$o = AQ + QA^* + \underline{H}[Q] + V, \tag{2.2}$$

$$\underline{H}[Q] \triangleq \underline{E}_{\alpha,w}[i\alpha(\lambda xx^* - xx^*\lambda)], \tag{2.3}$$

where we note that regardless of the statistics of α, $\underline{H}$ as defined above is a linear mapping from $\underline{R}^{nxn}$ into itself. Normalizing the random variable $\tilde{\alpha} = \alpha/\sigma$, it turns out that

$$\underline{H}_{KK}[Q] = 0 \qquad \text{a.} \quad (2.4)$$

$$\lim_{\sigma\uparrow\infty} \underline{H}_{KJ}[Q] = -\frac{1}{T}\,|\lambda_k - \lambda_j|\, Q_{KJ} \qquad \text{b.} \quad (2.4)$$

$$T \triangleq \frac{1}{\sigma}\int_0^\infty dt\underline{E}\ [\cos\tilde{\alpha}t] = \int_0^\infty dt\underline{E}\ [\cos\alpha t] \qquad \text{c.} \quad (2.4)$$

Equation (2.4.a) follows by definition and (2.4.b) governs the asymptotic behavior of $\underline{H}[Q]$ for large uncertainty levels. The quantity T, termed the relaxation time-scale, plays a key role in determining the magnitude of many phenomena associated with the impact of parameter uncertainties

Equations (2.4) summarize the required maximum data set for parameter statistics. Note that precisely this data, assumed available for the perturbation $i\alpha\lambda$, does not, in fact, imply either: (1) anything about the statistical dependence of the diagonal terms $i\alpha\lambda_K$ of $i\alpha\lambda$,; or (2) that $i\alpha\lambda_K$ is a random variable, constant in time, i.e., the data (2.4) permits each $i\alpha\lambda_K$ to be a stationary random process.

To be clear about what, under data (2.4), is known, as opposed to what is not known, we restate the situation as follows. The system evolves in time according to

$$\dot{x} = Ax + i\hat{\lambda}x + w, \qquad \text{a.}$$
$$\hat{\lambda} = \operatorname{diag}[\hat{\lambda}_K(t)], \qquad \text{b.} \tag{2.5}$$

where each $\hat{\lambda}_K$ is a real-valued stationary random process with zero mean and finite total power and $\hat{\lambda}$ is such that, with the normalization constant σ introduced as above, relations (2.4) hold for the quantity

$$\underline{H}[Q] \triangleq \underline{E}_{\lambda,w}[i(\hat{\lambda}xx^* - xx^*\hat{\lambda})].$$

To select a maximally unpresumptive stochastic model of $\hat{\lambda}$ out of the infinitely many models consistent with the available data, we now define a measure of a priori ignorance associated with the $\hat{\lambda}_K$'s.

First, let $\underline{T}_m$ denote some partition of the real line with $-\infty < t_o \leq t_1 \leq \ldots \leq t_m$, and define the random variables

$$\Lambda_K(t_i,t_{i+1}) = \int_{t_i}^{t_{i+1}} d\hat{\lambda}_K(t). \tag{2.6}$$

For a given $\underline{T}_m$, let $p(\Lambda;\underline{T}_m)$ denote the joint probability density of the $\Lambda_K(t_i,t_{i+1})$ for all k and i. Then, $\hat{\lambda}$ being a regular stationary process, the totality of the $p(\Lambda;\underline{T}_m)$ for all $\underline{T}_m$ and countable m uniquely defines the probability structure of $\hat{\lambda}$. Likewise, for given $\underline{T}_m$, the entropy

$$\underline{S}(\Lambda;\underline{T}_m) \triangleq -\int d\Lambda p(\Lambda;T_m)\ell n\, p(\Lambda;\underline{T}_m) \tag{2.7}$$

characterizes the degree of a priori ignorance with regard to the random variables $\Lambda_K(t_i, t_{i+1})$, K=1, ... n, i = 0, 1, ..., m, and the totality of the $\underline{S}(\Lambda;T_m)$ for all T_m describes the a priori ignorance of the process $\hat{\lambda}(t)$. We now seek a probability law for $\hat{\lambda}$ which maximizes $\underline{S}(\Lambda;T_m)$ for all T_m.

Now note that if the intervals $(t_i, t_{i+1}]$ and $(t_j, t_{j+1}]$ are disjoint, then by considering the joint statistics of $\Lambda_K(t_i, t_{i+1})$ and $\Lambda_K(t_j, t_{j+1})$, it follows from elementary properties of the entropy functional ([177]) that $S(\Lambda; \underline{T}_m)$ is maximized when $\Lambda_K(t_i, t_{i+1})$ and $\Lambda_K(t_j, t_{j+1})$ are statistically independent. Such independence does not violate the presumed data since (2.4) admits Λ_K's with independent increments. Thus, $\underline{S}(\Lambda; \underline{T}_m)$ is maximized when each $\int_0^t d\lambda_K(t)$ is a process with independent, stationary increments. However, this maximal value is not attained within the class of processes having finite total power; it is only attained as a supremal value over the class in the limit as statistical dependence between disjoint increments passes to zero. In [2] it was shown that the correct model for this limiting probability law is not Ito white noise but rather white noise under the Stratonovich interpretation of stochastic integrals.

Thus, considering joint statistics of disjoint increments of $\int_0^t d\lambda_K(t)$ for each K, the maximal (or rather supremal) value of $\underline{S}(\Lambda; \underline{T}_m)$ is attained (for all $\underline{T}_m$) when all of the λ_K's are Stratonovich white noise processes. This supremal value is approached without violating the constraints imposed by available data (2.4) and it remains only to determine the incremental covariance of the Weiner processes $\int_0^t d\lambda_K$, $K = 1, \ldots, N$. The data (2.4) does, in fact, afford a unique determination summarized as follows.

> For all $\underline{T}_m$, $\underline{S}(\Lambda; T_m)$ attains its supremal value within the constraints (2.4) when $X(t)$ evolves according to the Stratonovich white noise model.

$$dX_t = (A - \frac{1}{2}\{\rho\})\, X_t dt + i d\hat{\lambda} X_t + dW_t \tag{2.8}$$

where W_t is a Wiener process with intensity matrix V and the incremental covariance matrix, ρ, of the Weiner process $\hat{\lambda}$ is given by

$$\rho_{Kj} = \frac{1}{T}\,[\,|\lambda_K| + |\lambda_j| - |\lambda_K - \lambda_j|\,] \tag{2.9}$$

The above equation for X is interpreted as an Ito stochastic differential equation and the term $-\frac{1}{2}\{\rho\}$ is the so-called Stratonovich correction where we introduce the notation:

$$\{M\} \triangleq \text{diag}\,\{M_{11}, M_{22}, \ldots, M_{NN}\} \tag{2.10}$$

for any square matrix M.

The following properties of the above stochastic model are easily shown.

Lemma 2.1.1

Consider the stochastic system given by (2.8), (2.9) and accompanying definitions. Then

a. ρ is nonnegative definite. (Thus $\hat{\lambda}$ is a well-defined Wiener process.)

b. The steady-state second moment matrix satisfies

$$0 = (A - \frac{1}{2}\{\rho\})Q + Q(A - \frac{1}{2}\{\rho\})^* + \rho \oplus Q + V, \tag{2.11}$$

where $\oplus$ denotes the Hadamard (element-by-element) product of two matrices. Alternately, (2.11) can be written as

$$\begin{aligned} &0 = AQ + QA^* + \underline{H}[Q] + V && \text{a.} \\ &\underline{H}_{Kj}\,[Q] = -\frac{1}{T}\,|\lambda_K - \lambda_j|\; Q_{Kj} && \text{b.} \end{aligned} \tag{2.12}$$

(Thus, the constraints (2.4) are identically satisfied.)

2.2 Maximum Entropy Model Under Parameter Bounds

In this section we briefly outline the main results of [34]. Consider the linear system

$$\dot{x}_\alpha = (A + \alpha\lambda)X_\alpha + w, \tag{2.13}$$

where the notation "x_α" emphasizes the dependence of the state on the imperfectly known parameter α. Here we assume, as is often the case in practice, that the only available information concerning α is bounds, which for present purposes are of the form

$$\alpha \in [-\sigma, \sigma] \tag{2.14}$$

The equation for the second moment Q_α for x_α conditioned on $\alpha(t)$ has the form

$$\dot{Q}_\alpha = AQ_\alpha + Q_\alpha A^* + \underline{M}[Q_\alpha, \alpha] + V \tag{2.15}$$

The problem we pursue is the following. Assuming only bounds on the <u>stochastic modification term $\underline{M}$</u> in [2.15] of the form:

$$|\underline{M}\, [Q_\alpha, \alpha]_{Kj}\, | \leq \sigma |\ (\lambda_K + \lambda_j^*)\ Q_{\alpha Kj}|, \tag{2.16}$$

determine a realization of α which maximizes the <u>parameter entropy</u>.

<u>Lemma 2.2.1</u>

The induced maximum entropy model for the second moment Q subject to (2.16) is given by

$$\dot{Q} = AQ + QA^* + \underline{H}_Q\, [Q] + V \tag{2.17}$$
$$(H_Q\ \ [Q])_{Kj} \triangleq -\sigma |\lambda_K + \lambda_j^*|\ Q_{Kj}$$

The usefulness of Lemma 2.2.1 lies in the fact that it leads to the following <u>extended</u> optimization problem. Determine feedback compensation gains which minimize (interpret all notation as closed-loop; see [9])

$$J_e = \mathrm{tr}\ [Q\hat{R}] = \mathrm{tr}\ [P\hat{V}]$$

subject to

$$0 = AQ + QA^* + \underline{H}_Q\,[Q] + \hat{V},$$
$$0 = A^*P + PA + \underline{H}_P\,[P] + \hat{R},$$

where

$$\hat{V} \triangleq V + \underline{V}\,[Q],$$
$$\hat{R} \triangleq R + \underline{R}\,[P],$$
$$\underline{V}\,[Q] \triangleq \operatorname*{diag}_K \left\{ \sum_{\ell} 2\,\sigma\,|\lambda_K + \lambda_\ell^*|\;|Q_{K\ell}| \right\}$$
$$\underline{R}\,[P] \triangleq \operatorname*{diag}_K \left\{ \sum_{\ell} 2\,\sigma\,|\lambda_K^* + \lambda_\ell|\;|P_{K\ell}| \right\}$$

The main result, which follows, shows that under the formulation of the extended problem, the cost can be bounded from above (i.e., guaranteed) over the assumed parameter range.

Theorem 2.2.1

If $\overline{Q}$, $\overline{P}$ solve the extended problem, then

$$\overline{Q} \geq Q,\; \overline{P} \geq P,$$

for all P and Q solving the original problem, and

$$J_e\,(\overline{Q}, \overline{P}) \geq J_e(\alpha).$$

Furthermore, $A + \alpha\lambda$ is stable for all $\alpha \epsilon [-\sigma, \sigma]$.

The details of the above formulation together with the proof of the above theorem are to appear in [34]. For the present, we note that this formulation opens the way to a stochastic treatment which guarantees stability over the stipulated ranges of parameter variations.

SECTION 3.0

THE OPTIMAL PROJECTION APPROACH TO REDUCED-ORDER DESIGN

3.0 THE OPTIMAL PROJECTION APPROACH TO REDUCED ORDER DESIGN

3.1 Review of Optimal Projection Results

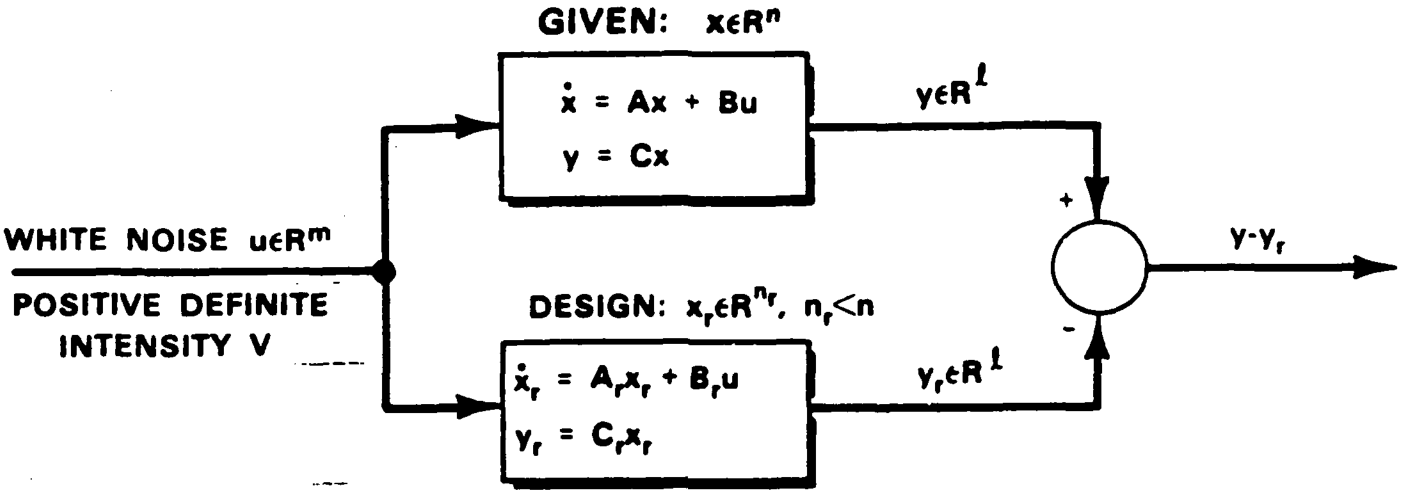
Optimal projection design is based on a series of three results which characterize the quadratically optimal reduced-order model, reduced-order state estimator and reduced-order dynamic compensator. Assuming a purely dynamic linear structure for the desired system (model, estimator, or compensator), whose order is determined by implementation constraints, a parameter optimization approach is taken. There is, of course, nothing novel about this approach per se and it has been widely studied in the model reduction and control literature ([38,40,46,72-98]). This approach, however, fell into disrepute because of the extreme complexity of the grossly unwieldy first-order necessary conditions which afforded little insight and engendered brute-force gradient search techniques. The crucial discovery occurred in [7] where it was revealed that the necessary conditions for the dynamic-compensation problem give rise to the definition of an optimal projection as a rigorous, unassailable consequence of quadratic optimality without recourse to *ad hoc* methods as in [99-102]. Exploitation of this projection leads to immense simplification of the "primitive" form of the necessary conditions for each of the three problems. As summarized in Figure 3.1-1, the modelling, estimation, and compensation design equations form a natural progression: coupled systems of two, three, and four matrix equations whose solutions determine the desired gains (A_m, B_m, C_m), (A_e, B_e, C_e), and (A_c, B_c, C_c). The novel equations are the modified Lyapunov equations for the reduced-order modelling problem, versions of which arise in the estimation and compensation problems. Since the modified Riccati equations encompass the standard observer and regulator Riccati equations, the optimal projection equations for the reduced-order state-estimation and dynamic-compensation problems provide a fundamental generalization of steady-state Kalman filter and LQG theory.

3.1.1 Optimal Model Reduction

The optimal model-reduction problem (Figure 3.1.1-1) involves determining a low-order model that minimizes the steady-state, quadratically weighted output error when the original system and reduced-order model are subjected to white noise inputs. This problem formulation is particularly appropriate when the reduced-order model is used to simulate the statistical

DESIGN PROBLEM	DYNAMICAL SYSTEM	DESIGN OBJECTIVE	PERFORMANCE CRITERION	DESIGN EQUATIONS
REDUCED-ORDER MODELLING	$\dot{x} = Ax + Bw$ $y = Cx$	$\dot{x}_m = A_m x_m + B_m w$ $y_m = C_m x_m$	$\lim_{t\to\infty} E(y-y_m)^T R(y-y_m)$	2 Lyapunov
REDUCED-ORDER STATE ESTIMATION	$\dot{x} = Ax + w_1$ $y = Cx + w_2$	$\dot{x}_e = A_e x_e + B_e y$ $y_e = C_e x_e$	$\lim_{t\to\infty} E(Lx-y_e)^T R(Lx-y_e)$	1 Riccati 2 Lyapunov
REDUCED-ORDER DYNAMIC COMPENSATION	$\dot{x} = Ax + Bu + w_1$ $y = Cx + w_2$	$\dot{x}_c = A_c x_c + B_c y$ $u = C_c x_c$	$\lim_{t\to\infty} E(x^T R_1 x + u^T R_2 u)$	2 Riccati 2 Lyapunov

Figure 3.1-1. Optimal Projection Design



STEADY STATE TRACKING CRITERION

$$J(A_r, B_r, C_r) \equiv \lim_{t \to \infty} E[(y-y_r)^T R(y-y_r)] \quad \text{(R positive definite)}$$

ASSUME: A, A_r STABLE
(A_r, B_r, C_r) MINIMAL

Figure 3.1.1-1. Optimal Model-Reduction Problem

response of the high-order system; no claim whatsoever is made as to its usefulness for estimator or controller design. The main result (Figures 3.1.1-2 and 3.1.1-3) involves a coupled system of two nxn modified Lyapunov equations whose solutions are given by a pair of rank-n_r controllability and observability pseudogramians $\hat{Q}$ and $\hat{P}$. The matrix $\mathcal{T}$ coupling these equations is idempotent since

$$\mathcal{T}^2 = G^T\Gamma G^T\Gamma = G^T I_n \Gamma = \mathcal{T}.$$

This oblique projection determines the optimal reduced-order model via an aggregation as a direct consequence of optimality.

Since the optimal projection equations for model reduction are first-order necessary conditions for optimality, they may possess nonunique solutions corresponding to multiple local extrema (Figure 3.1.1-4). The mechanism responsible for this phenomenon becomes clear upon characterizing the optimal projection as a sum of rank-1 eigenprojections of the product of the solutions of an equivalent system of "standard" Lyapunov equations (Figures 3.1.1-5 and 3.1.1-6): The first-order necessary conditions are ambiguous in the sense that they fail to specify which n_r eigenprojections comprise the optimal projection corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem. Specifically, since the pseudogramians $\hat{Q}$ and $\hat{P}$ can be rank deficient in $\binom{n}{n_r} = \frac{n!}{n_r!(n-n_r)!}$ ways, there may be precisely this many "extremal" projections corresponding to an identical number of local extrema.

An immediate consequence of this insight is a rigorous extremality context for Moore's balancing method ([50]) thereby demonstrating its quadratic nonoptimality. Specifically, examples can easily be constructed for which the reduced-order model obtained from the balancing method is much worse with respect to the least-squares criterion than the quadratically optimal reduced-order model (Figure 3.1.1-7). In general, all that can be said is that the presence of a weak subsystem (in the sense of Moore) indicates that the reduced-order model obtained by truncation in the balanced basis may be in the proximity of an extremal of the optimal model-reduction problem; however, this extremal may very well be a local (or even global) maximum.

LEMMA. IF $\hat{Q}$ AND $\hat{P}$ ARE NONNEGATIVE DEFINITE THEN THE PRODUCT $\hat{Q}\hat{P}$ IS NONNEGATIVE SEMISIMPLE. HENCE IF RANK $\hat{Q}\hat{P} = n_r$ THEN THERE EXIST $G, \Gamma \epsilon R^{n_r x n}$ AND POSITIVE SEMISIMPLE $M \epsilon R^{n_r x n_r}$ SUCH THAT

$$\hat{Q}\hat{P} = G^T M \Gamma$$

$$\Gamma G^T = I_{n_r}$$

Figure 3.1.1-2. Factorization Lemma

IF ADMISSIBLE (A_r, B_r, C_r) IS OPTIMAL THEN THERE EXIST $n \times n$ NONNEGATIVE-DEFINITE MATRICES $\hat{Q}$ AND $\hat{P}$ SUCH THAT, FOR SOME (G, M, Γ)-FACTORIZATION OF $\hat{Q}\hat{P}$, (A_r, B_r, C_r) ARE GIVEN BY

$$A_r = \Gamma A G^T,$$

$$B_r = \Gamma B,$$

$$C_r = C G^T,$$

AND SUCH THAT, WITH $\tau \equiv G^T \Gamma$ AND $\tau_\perp \equiv I_n - \tau$,

$$0 = A\hat{Q} + \hat{Q}A^T + BVB^T - \tau_\perp BVB^T \tau_\perp^T,$$

$$0 = A^T\hat{P} + \hat{P}A + C^TRC - \tau_\perp^T C^TRC\tau_\perp,$$

$$\text{RANK } \hat{Q} = \text{RANK } \hat{P} = \text{RANK } \hat{Q}\hat{P} = n_r.$$

Figure 3.1.1-3. Optimal Reduced-Order Model

- $A = \mathrm{diag}(-\alpha_1, \ldots, -\alpha_n),\ \alpha_i > 0$
  - $BB^T = \mathrm{diag}(\beta_1, \ldots, \beta_n),\ C^TC = \mathrm{diag}(\gamma_1, \ldots, \gamma_n),\ \beta_i, \gamma_i > 0$
- $\widehat{Q} = \mathrm{diag}(\widehat{Q}_1, \ldots, \widehat{Q}_n),\ \widehat{P} = \mathrm{diag}(\widehat{P}_1, \ldots, \widehat{P}_n)$

$$\Rightarrow \widehat{Q}_i = \frac{\beta_i}{2\alpha_i}\delta_i\ ,\ \widehat{P}_i = \frac{\gamma_i}{2\alpha_i}\delta_i\ ,\ \delta_i = 0 \text{ or } 1$$

- But: Only n_r of the δ_i's are 1

$$\Leftrightarrow \binom{n}{n_r} = \frac{n!}{n_r!\,(n-n_r)!} \quad \text{possible extrema}$$

Figure 3.1.1-4. Existence of Multiple Extrema in Optimal Model Reduction

- $\hat{\hat{Q}}, \hat{\hat{P}}$ **nonnegative definite** $\Rightarrow \hat{\hat{Q}}\hat{\hat{P}}$ **nonnegative semisimple**

$$\bullet\ \hat{\hat{Q}}\hat{\hat{P}} = S\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}S^{-1}, \qquad \lambda_i \geq 0$$

$$= S\left(\sum_{i=1}^{n} \lambda_i E_i\right)S^{-1}, \qquad E_i = \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0\,1\,0 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \Pi_i[\hat{\hat{Q}}\hat{\hat{P}}], \qquad \Pi_i[\hat{\hat{Q}}\hat{\hat{P}}] = SE_iS^{-1}$$

- $\Rightarrow\ \tau = \sum_{j=1}^{n_r} \Pi_{i_j}[\hat{\hat{Q}}\hat{\hat{P}}]$

Figure 3.1.1-5. Obtaining an Oblique Projection From an Eigenprojection Decomposition of a Semisimple Matrix

$$0 = (A - \tau A \tau_{\perp})\hat{\hat{Q}} + \hat{\hat{Q}}(A - \tau A \tau_{\perp})^T + BVB^T$$

$$0 = (A - \tau_{\perp} A \tau)^T \hat{\hat{P}} + \hat{\hat{P}}(A - \tau_{\perp} A \tau) + C^T R C$$

$$\tau = \sum_{r=1}^{n_r} \Pi_{i_r}[QP]$$

Figure 3.1.1-6. Optimal Projection Equations for Model Reduction: Standard Lyapunov Equation Form

- $n = 2$, $\alpha_1 = 1$, $\alpha_2 = 10^6$, $\beta_1 = 1$, $\beta_2 = 10^6$,

  $\gamma_1 = 1$, $\gamma_2 = 10^3$

  $\Rightarrow$ second-order modes $\sigma_1 = .5$, $\sigma_2 \approx .012$

- Moore's approach $\Rightarrow$ truncate $x_2 \Rightarrow J = 500$

  However: truncate $x_1 \Rightarrow J = .5$

- This attainment of a global maximum is a direct result of incorrect selection of n_r eigenprojections in constructing the optimal projection

Figure 3.1.1-7. Inadequacy of the Singular Values in Quadratically Optimal Model Reduction

Although the optimal projection equations characterize all extrema, it is desirable for theoretical and practical reasons to directly characterize the global minimum. One approach, investigated in [29], involves decomposing the cost as

$$J = \sum_{i=1}^{n} J_i,$$

where each J_i corresponds to a particular eigenprojection. This technique, which is reminiscent of Skelton's Component Cost Analysis [52,101], permits rapid sorting of the local extrema by directing the algorithm to the global optimum (Figure 3.1.1-8).

3.1.2 Optimal Reduced-Order Dynamic Compensation

Virtually all research into the design of reduced-order controllers involves one of two sequential procedures: model reduction followed by controller design or controller design followed by controller reduction (Figure 3.1.2-1). The optimal projection equations represent a radical departure from both of these approaches by directly characterizing the quadratically optimal reduced-order controller for a high-order model. The form of the necessary conditions (a coupled system of two modified Riccati equations and two modified Lyapunov equations given by equations (2.18)-(2.21) of [23] in Appendix B) indicates the essential presence of the LQG and model-reduction operations. The coupling by means of the projection, however, reveals the inherent inseparability of these operations in the reduced-order case and represents a graphic portrayal of the demise of the classical separation principle. Hence, optimality considerations demand that, in a very precise sense, reduction and control design be performed simultaneously.

The optimal projection equations also reveal the suboptimality of ad hoc controller reduction methods. As shown in [19] (Appendix A), several reduction methods involve equations similar to the optimal projection equations but lack crucial coupling terms. Prior to the discovery of the optimal projection equations, the state of affairs in reduced-order controller design was philosophically analogous to fluid mechanics should it have developed without benefit of the Navier-Stokes equations: Instead of deciding which of the numerous

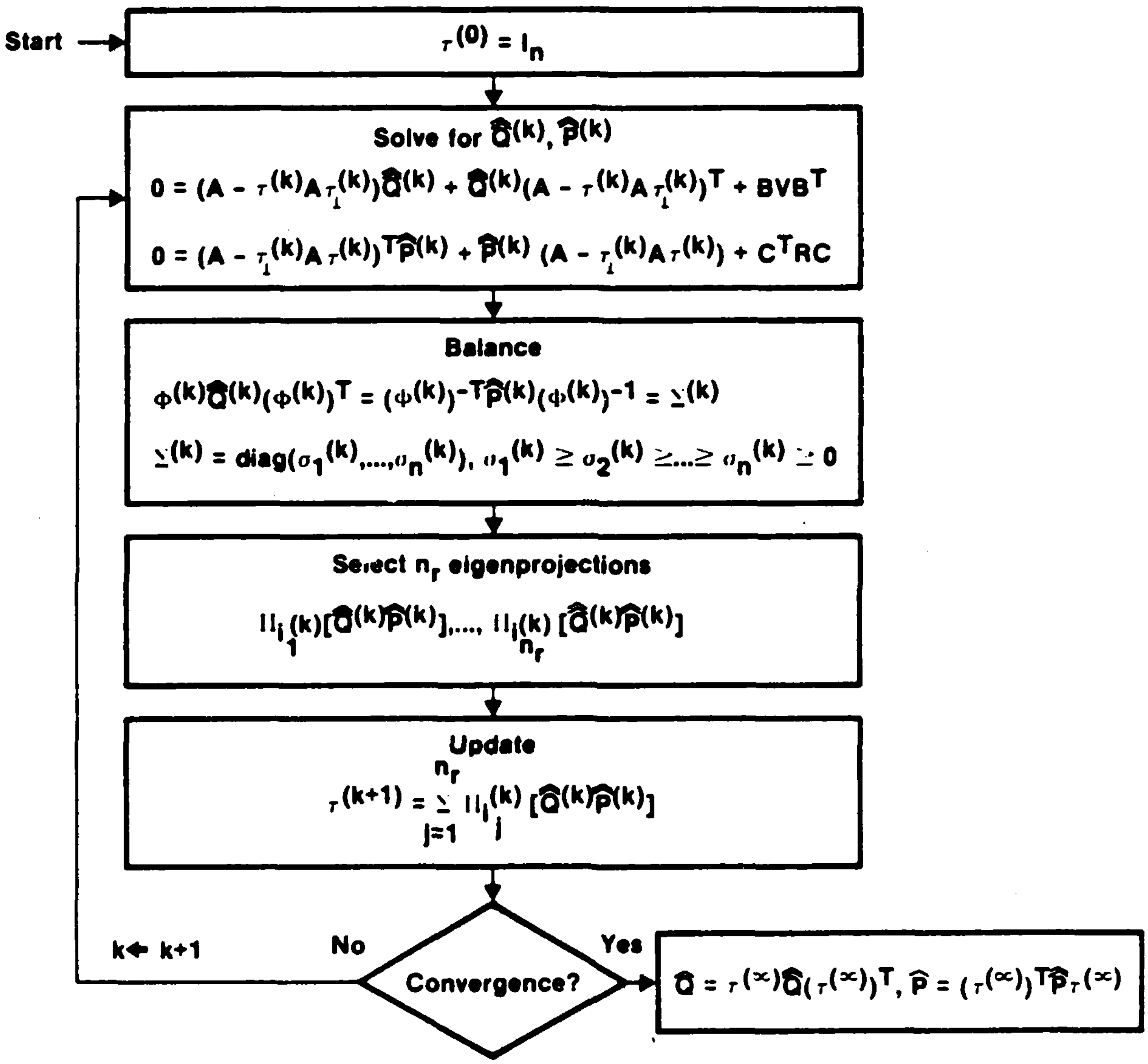


Figure 3.1.1-8. Computational Algorithm for Optimal Model Reduction

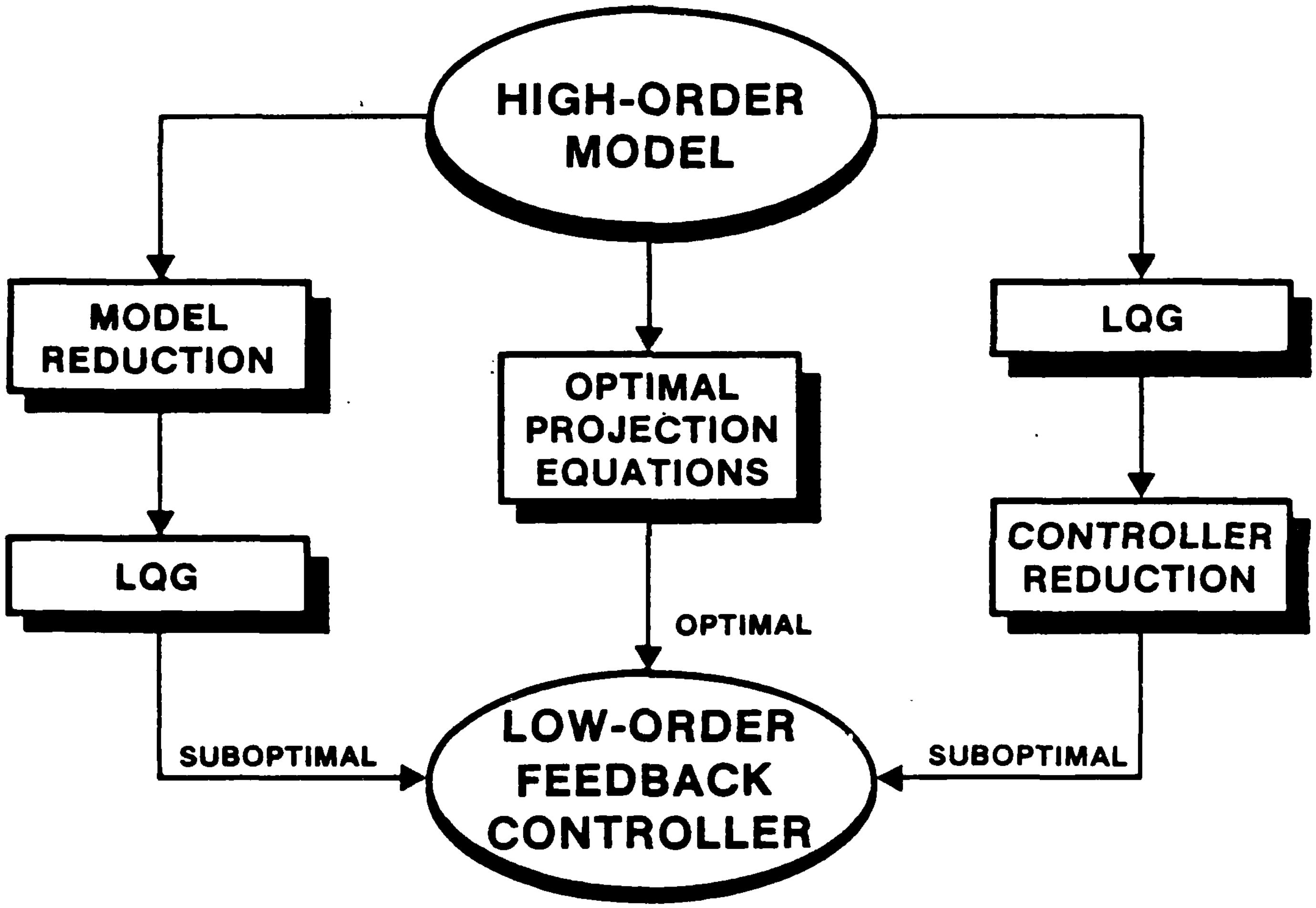


Figure 3.1.2-1. Optimal Projection Design for Reduced-Order Dynamic Compensation

terms are small and hence can be neglected, engineers would have been obliged to laboriously construct the important terms!

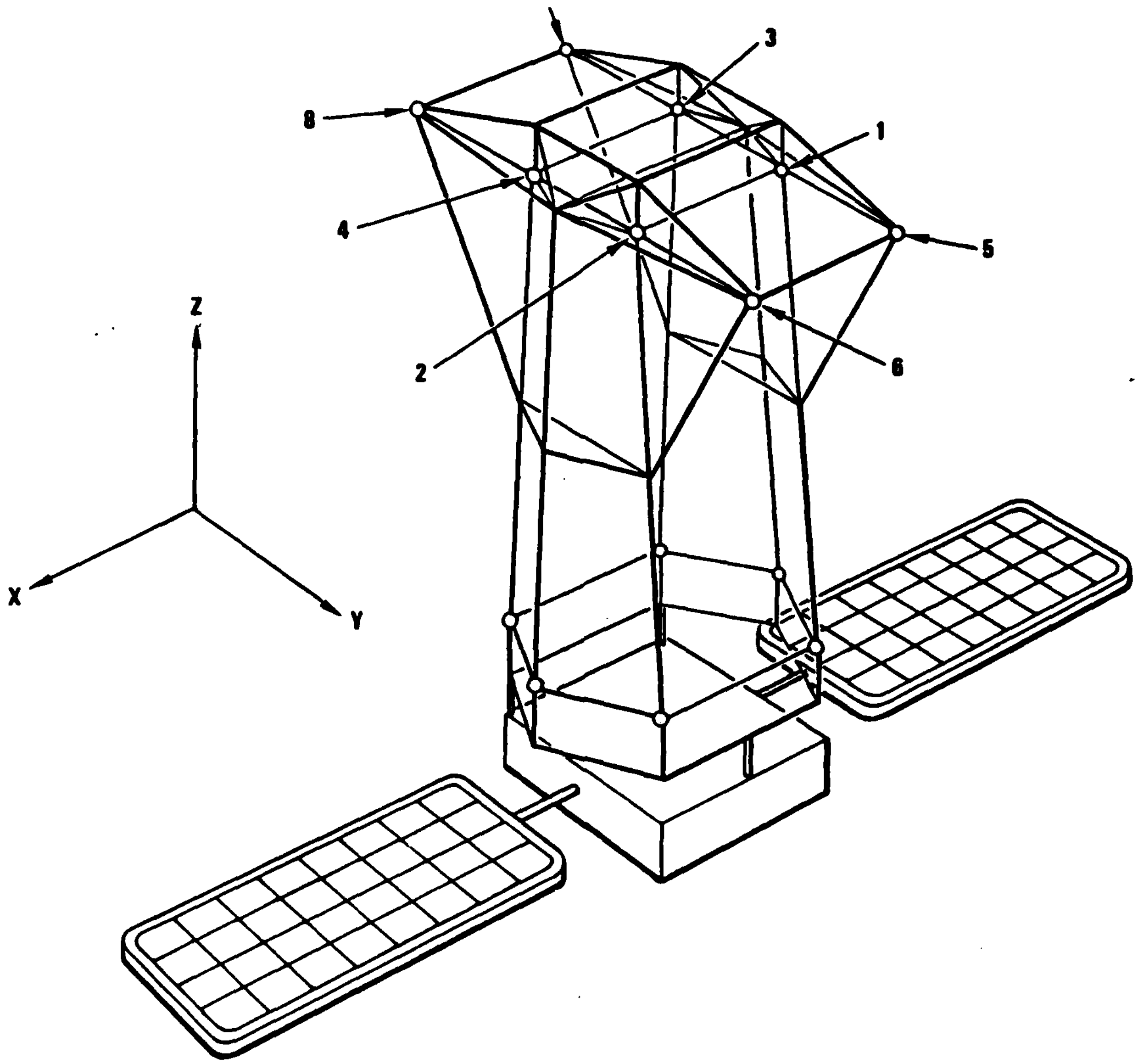
3.1.2.1 Optimal Projection Controller Design for Draper Model #2

The optimal projection approach was applied to a 20-state version of the CSDL Model #2 (see Figure 3.1.2.1-1 [19]) excluding modelling uncertainties. This corresponds to Task 2.A mentioned in Section 1.0, i.e., the "high authority" control design.

This example was used to compare both theoretically and numerically the optimal projection approach with a variety of suboptimal controller-order reduction methods. The theoretical comparison shows that all current suboptimal techniques essentially define a (suboptimal) projection characterizing the reduced-order compensator. In contrast, the optimal projection design equations define the needed projection by rigorous application of optimality principles. Moreover, all the approaches considered in [19] can be displayed in a common notation, and this graphically reveals the suboptimal design equations as special cases of or approximations to the optimal projection equations.

For numerical comparison it is standard procedure to plot the regulation cost $E[x^T R_1 x]$ as a function of control cost $E[u^T R_2 u]$. Results for these tradeoff curves are shown in Figure 3.1.2.1-2. The very bottommost curve represents the full-order, LQG design. Since this is the best obtainable when there is no restriction on compensator order, the problem is obtaining a lower order design whose tradeoff curve is as close to the LQG results as possible.

The thin black lines in Figure 3.1.2.1-2 show the $n_c = 10$, 6, and 4 designs obtained via Component Cost Analysis [101], where n_c denotes the compensator dimension. This appears to be the most successful suboptimal method applied to the example problem considered here. Note that the 10th and 6th order compensator designs are quite good, but when compensator order is sufficiently low ($n_c = 4$) and controller bandwidth sufficiently large ($R<5.0$), the method fails to yield stable designs. This difficulty is characteristic of all suboptimal techniques surveyed, and, in fairness, it should be noted that most other suboptimal design methods fail to give stable designs for compensator orders below 10.



REFERENCE:

R. E. Skelton and P. C. Hughes, "Modal Cost Analysis for Linear Matrix Second-Order Systems," J Dyn Syst Meas and Control Vol 102, September 1980, pp 151-180

13466-3

Figure 3.1.2.1-1. CSDL Model #2

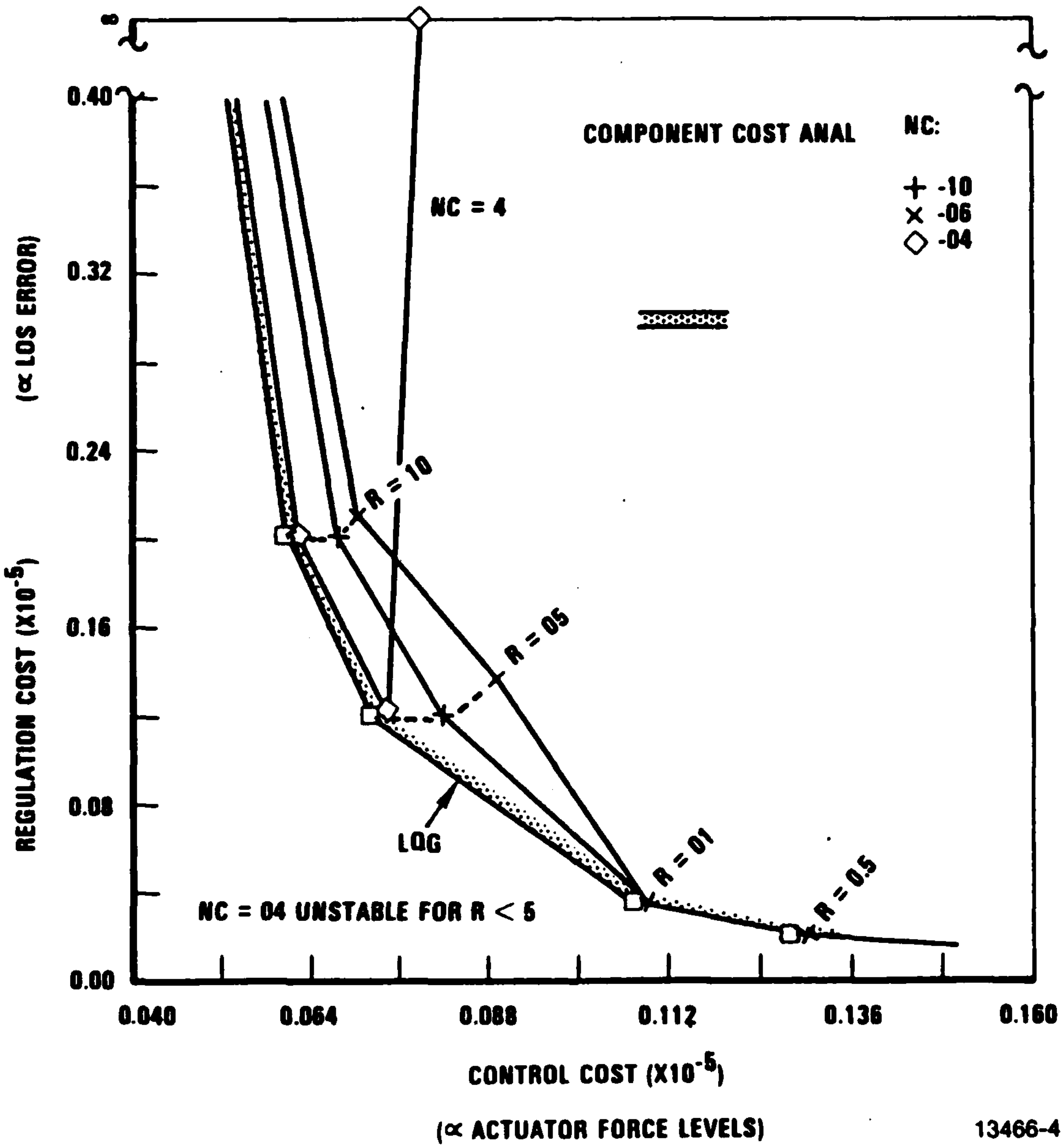


Figure 3.1.2.1-2. Cost Tradeoff Curves

In contrast, the width of the grey line in Figure 3.1.2.1-2 encompasses all the optimal projection results for compensators of orders 10, 6, and 4.

3.1.3 Review of Infinite-Dimensional Results

The optimal projection equations for fixed-order dynamic compensation have been generalized in [28] (Appendix D) to the case in which the controlled system is an infinite-dimensional system in Hilbert space. Mathematically, the dynamics operator A is assumed to generate a C_o semigroup, the stochastic differential equation is treated as an evolution equation, white noise is interpreted in the sense of Balakrishnan ([134]) and the input and output operators B and C are assumed to be bounded.

In addition to the above mathematical considerations associated with the physical description of the distributed parameter system, optimal projection design directly addresses the practical constraints according to Athans ([130]) and Balas ([131]) of: 1) finitely many sensors and actuators, 2) a finite-dimensional, controller, and 3) natural system dissipation. The validity of 2) is apparent from the fact that processing and transmitting electrical signals by conventional analog or digital components constitutes finite-dimensional action. Hence, although distributed parameter systems are most accurately represented by infinite-dimensional models, real-world considerations demand that implementable controllers be modelled as finite-dimensional lumped parameter systems.

Clearly, the above observations effectively preclude the possibility of realizing infinite-dimensional controllers that involve full-state feedback or full-state estimation (see, e.g., [132-134] and the numerous references therein). Although the transition to finite-dimensional, implementable controllers proceeds by means of approximation schemes, these results only guarantee optimality in the limit, i.e., as the order of the approximating controller increases without bound ([135-138]). Hence, there is no guarantee that a particular approximate (i.e., discretized) controller is actually optimal over the class of approximate controllers of a given order dictated by implementation constraints. Moreover, even if an optimal approximate finite-dimensional controller could be obtained, it would almost certainly be suboptimal in the class of all controllers of the given order.

Although the sequence of operations described above corresponds to the left-hand branch of Figure 3.1.3-1, one can alternatively follow the right-hand branch, i.e., replace the distributed parameter system with a high-order finite-dimensional model and utilize the optimal projection equations of [23] (Appendix B) to obtain a fixed-order controller. The most direct route, however, involves a direct characterization of the optimal finite-dimensional, fixed-order controller for the <u>original</u> distributed parameter system. The resulting equations are exactly analogous to the optimal projection equations obtained in [23] for the finite-dimensional case. Instead of a system of four matrix equations, however, the result now involves a system of four <u>operator</u> equations whose solutions characterize the optimal finite-dimensional fixed-order dynamic compensator (see (3.9)-(3.18) of [28] in Appendix D). Moreover, the optimal projection now becomes a bounded idempotent Hilbert-space operator whose rank is precisely equal to the order of the compensator. To illustrate the beauty and simplicity of this result, note, for example, that the $n_c x n_c$ dynamics matrix A_c of the controller is given by (see (3.9) of [28])

$$A_c = \Gamma(A - QBR_2^{-1}B^* - C^*V_2^{-1}CP)G^*$$

where G^* is a mapping from R^{n_c} into the domain of A and Γ is a mapping from the Hilbert space into R^{n_c}. Hence, the above expression is indeed a valid representation of an $n_c x n_c$ matrix which, most interestingly, incorporates an internal model of the full dynamics operator of the infinite-dimensional system!

Since the only explicit assumption on the unbounded dynamics operator is that it generate a strongly continuous semigroup, these results are potentially applicable to a broad range of specific partial and functional differential equations. Their <u>actual</u> applicability is essentially limited by practical constraint 3). Because of the steady-state problem setting, it is implicitly assumed that the distributed parameter system is stabilizable, i.e., that there exists a dynamic compensator of a given order such that the closed-loop system is uniformly stable. The stabilization problem has been considered in [142-148] for delay, parabolic, and damped hyperbolic systems.

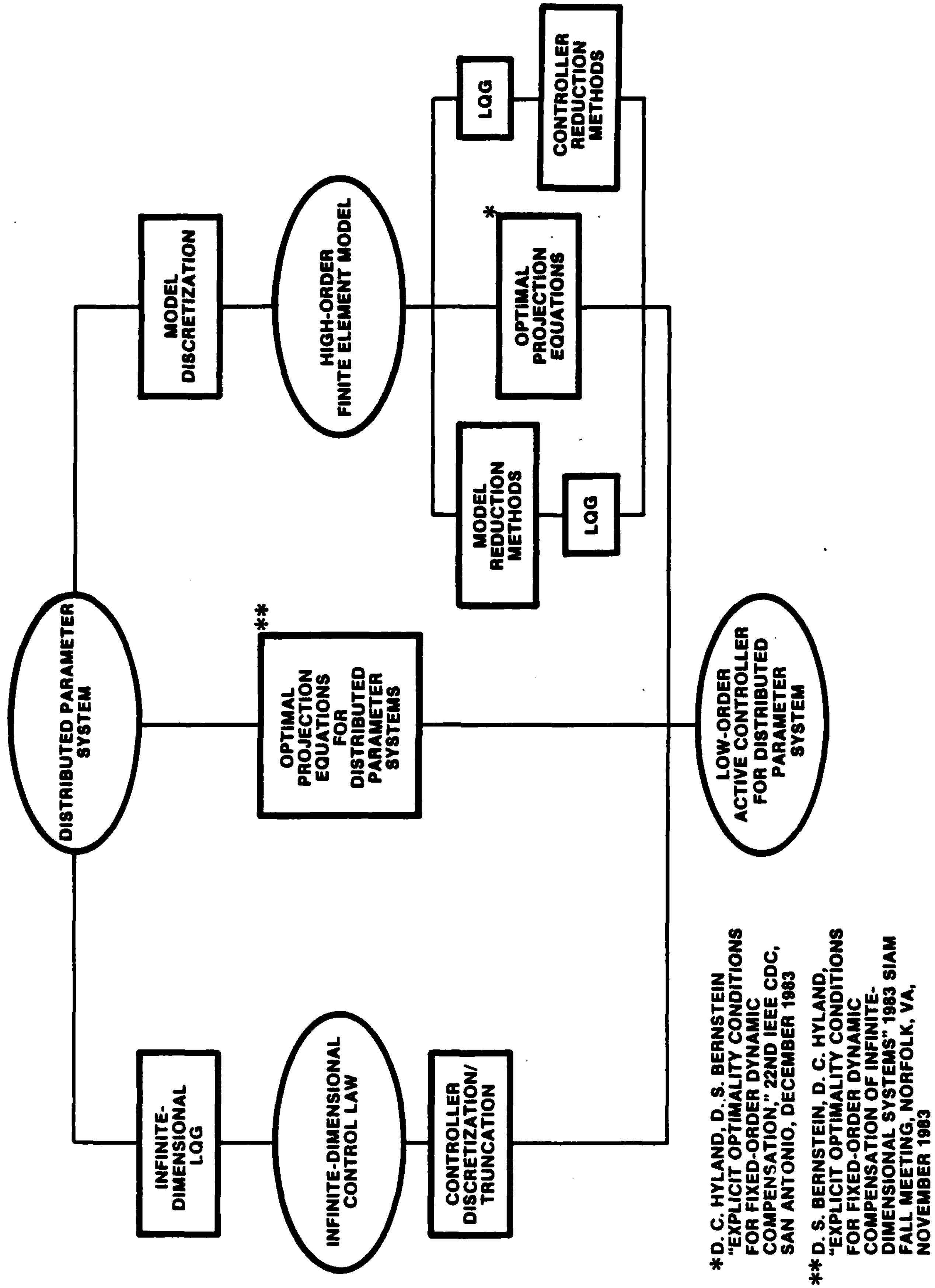


Figure 3.1.3-1. Optimal Projection Approach to Finite-Dimensional Fixed-Order Dynamic Compensation of Distributed Parameter Systems

SECTION 4.0

OPTIMAL PROJECTION/MAXIMUM ENTROPY DESIGN SYNTHESIS

4.0 OPTIMAL PROJECTION/MAXIMUM ENTROPY DESIGN SYNTHESIS

4.1 Design Equations

By combining maximum entropy stochastic modelling with optimal projection design, we obtain a powerful system design methodology which generalizes LQG theory in two fundamental respects: design of reduced-order controllers plus accommodation of *a priori* parameter uncertainties. The most general results obtained thus far apply to the modelling, estimation, and control problems and are presented in detail in [31] (see Appendix E). For the control problem, these results are summarized in Figures 4.1-1 to 4.1-4.

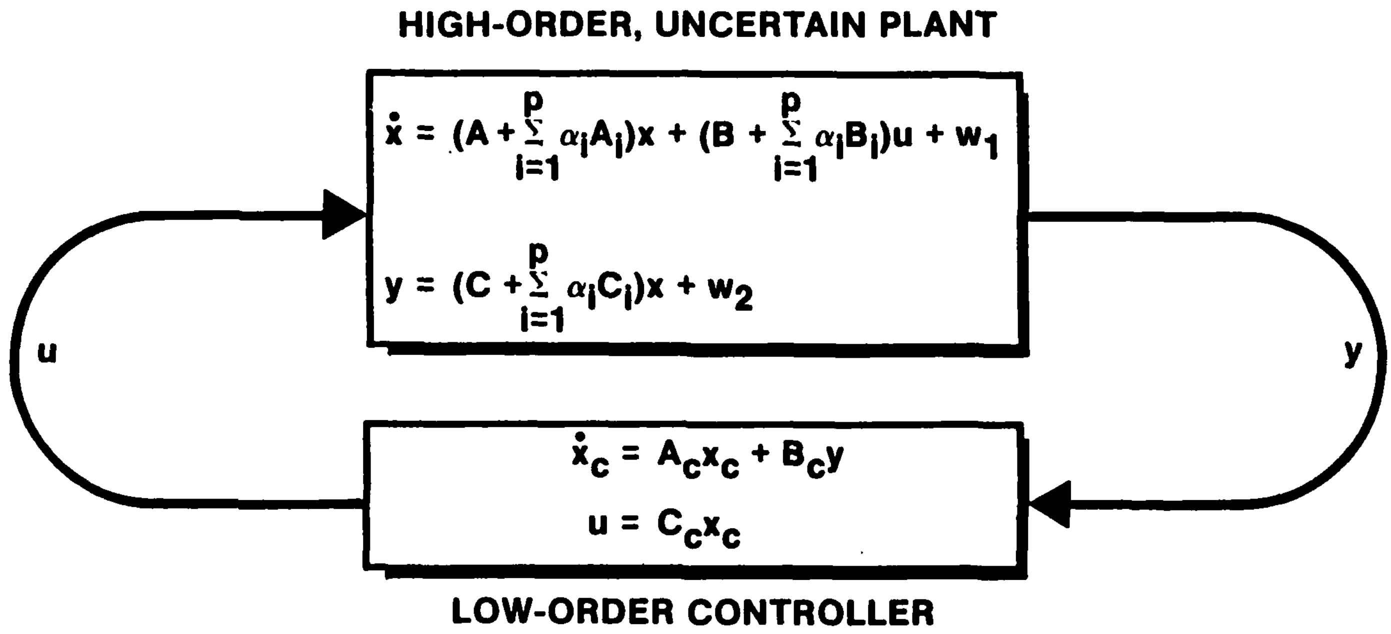
The control-design problem is summarized in Figure 4.1-1 and reveals the requirement of a low-order controller of fixed dimension and the technique of using Stratonovich white noise to model parameter uncertainties. Figure 4.1-2 summarizes the stability conditions under which the optimization is carried out. The optimal controller gains A_c, B_c and C_c are given in Figure 4.1-3 in terms of Q, P, $\hat{Q}$ and $\hat{P}$ which are, in turn, determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations (shown in Figure 4.1-4).

4.2 Combined OP/ME Design for the NASA SCOLE Model

Harris GASD recently completed a NASA/LaRC supported study on the Spacecraft Control Laboratory Experiment (SCOLE) configuration shown in Figure 4.2-1 which is the subject of the NASA/IEEE Design Challenge. Full details of our model and design results are given in [27].

A high-order finite element model was constructed for SCOLE, treating the Shuttle and reflector as rigid bodies and the connecting mast as a classical beam with torsional stiffness. This model includes the Shuttle products-of-inertia and the offset between reflector center-of-mass and its attachment point on the mast. The quadratic performance penalty on the system state is simply the total mean square line-of-sight error.

As part of the study, we considered a system model including the first eight modes and (1) performed LQC studies to select the control authority and establish a baseline and (2) designed full-order (16-state) compensators with a



PERFORMANCE CRITERION

$$J(A_c, B_c, C_c) = \lim_{t \to \infty} E[x^T R_1 x + 2x^T R_{12} u + u^T R_2 u]$$

Technical Assumption: $B_i \neq 0 \Rightarrow C_i = 0$

Figure 4.1-1. Steady-State Reduced-Order Dynamic-Compensation Problem with Parameter Uncertainties

Nominal Closed-Loop Dynamics Matrix

$$\tilde{A} = \begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix}$$

Uncertainty Due to i^th Uncertain Parameter

$$\tilde{A}_i = \begin{bmatrix} A_i & B_iC_c \\ B_cC_i & 0 \end{bmatrix}$$

Corrected Dynamics Matrix

$$\tilde{A}_s = \tilde{A} + \frac{1}{2}\sum_{i=1}^{p} A_i^2$$

STABILITY IS DETERMINED BY

$$A_s \otimes I_{n+n_c} + I_{n+n_c} \otimes A_s + \sum_{i=1}^{p} A_i \otimes A_i$$

$\otimes$ = Kronecker product

Figure 4.1-2. Second-Moment Stability

CONTROLLER GAINS (Functions of Q, P, $\hat{Q}$, $\hat{P}$)

$$A_c = \Gamma(A_s - B_s R_{2s}^{-1}\mathcal{P}_s - \mathcal{Q}_s V_{2s}^{-1} C_s)G^T$$

$$B_c = \Gamma \mathcal{Q}_s V_{2s}^{-1}$$

$$C_c = -R_{2s}^{-1}\mathcal{P}_s G^T$$

NOTATION

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad \Gamma G^T = I_{n_c} \; (\Rightarrow \tau = G^T \Gamma = \tau^2)$$

$$\mathcal{A}Q\mathcal{A}^T = \sum_{i=1}^{p} A_i Q A_i^T, \quad \mathcal{A}Q\mathcal{B} = \sum_{i=1}^{p} A_i Q B_i, \text{ etc.}$$

$$A_s = A + \frac{1}{2}\mathcal{A}^2 \qquad B_s = B + \frac{1}{2}\mathcal{A}\mathcal{B} \qquad C_s = C + \frac{1}{2}\mathcal{C}\mathcal{A}$$

$$R_{2s} = R_2 + \mathcal{B}^T(P+\hat{P})\mathcal{B} \qquad V_{2s} = V_2 + \mathcal{C}(Q+\hat{Q})\mathcal{C}^T$$

$$\mathcal{Q}_s = QC_s^T + V_{12} + \mathcal{A}(Q+\hat{Q})\mathcal{C}^T \qquad \mathcal{P}_s = B_s^T P + R_{12}^T + \mathcal{B}^T(P+\hat{P})\mathcal{A}$$

Figure 4.1-3. Controller Gains

SOLVE FOR NONNEGATIVE-DEFINITE Q, P, $\hat{Q}$, $\hat{P}$

$$0 = A_sQ + QA_s^T + \mathcal{A}Q\mathcal{A}^T + V_1 + (\mathcal{A}-\mathcal{B}R_{2s}^{-1}\mathcal{P}_s)\hat{Q}(\mathcal{A}-\mathcal{B}R_{2s}^{-1}\mathcal{P}_s)^T - \mathcal{Q}_sV_{2s}^{-1}\mathcal{Q}_s^T + \tau_\perp\mathcal{Q}_sV_{2s}^{-1}\mathcal{Q}_s^T\tau_\perp^T$$

$$0 = A_s^TP + PA_s + \mathcal{A}^TP\mathcal{A} + R_1 + (\mathcal{A}-\mathcal{Q}_sV_{2s}^{-1}\mathcal{C})^T\hat{P}(\mathcal{A}-\mathcal{Q}_sV_{2s}^{-1}\mathcal{C}) - \mathcal{P}_s^TR_{2s}^{-1}\mathcal{P}_s + \tau_\perp^T\mathcal{P}_s^TR_{2s}^{-1}\mathcal{P}_s\tau_\perp$$

$$0 = (A_s-B_sR_{2s}^{-1}\mathcal{P}_s)\hat{Q} + \hat{Q}(A_s-B_sR_{2s}^{-1}\mathcal{P}_s)^T + \mathcal{Q}_sV_{2s}^{-1}\mathcal{Q}_s^T - \tau_\perp\mathcal{Q}_sV_{2s}^{-1}\mathcal{Q}_s^T\tau_\perp^T$$

$$0 = (A_s-\mathcal{Q}_sV_{2s}^{-1}C_s)^T\hat{P} + \hat{P}(A_s-\mathcal{Q}_sV_{2s}^{-1}C_s) + \mathcal{P}_s^TR_{2s}^{-1}\mathcal{P}_s - \tau_\perp^T\mathcal{P}_s^TR_{2s}^{-1}\mathcal{P}_s\tau_\perp$$

$$\text{RANK } \hat{Q} = \text{RANK } \hat{P} = \text{RANK } \hat{Q}\hat{P} = n_c$$

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} \qquad \tau_\perp = I_n - \tau$$

 # ⇨ GROUP GENERALIZED INVERSE

Figure 4.1-4. Optimal Projection/Maximum Entropy Design Equations

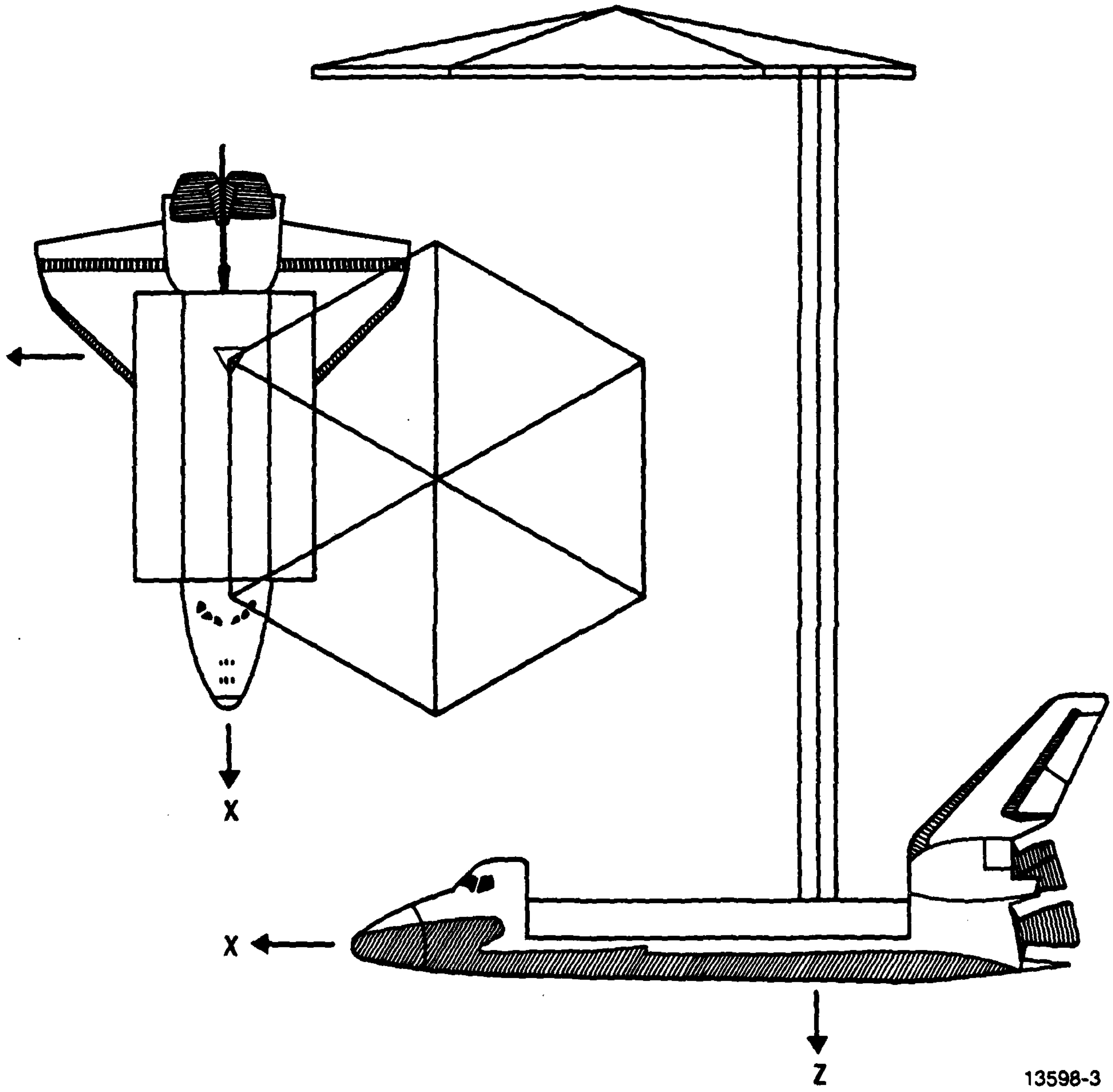


Figure 4.2-1. Spacecraft Control Lab Experiment (SCOLE)

maximum entropy model of modal frequency uncertainties. The maximum entropy model assumed that all elastic mode frequencies were subjected to independent variations (due to modelling error) of $+\sigma$ to $-\sigma$ relative to their nominal values. Thus the positive number σ denotes the overall fractional uncertainty.

Although robust stability is obtained under these independent and simultaneous variations, the robustness properties of specific designs are simply illustrated here by looking at the variation of performance and closed-loop poles when all modal frequencies are varied by the same fractional change from the nominal values. In other words, we interconnect a given controller design (be it LQG or maximum entropy) with a perturbed plant model wherein all modal frequencies are changed by δx (nominal values) and evaluate the closed-loop performance and pole locations. This is repeated for a range of values of δ.

Figure 4.2-2 shows how the pole locations for an LQG design wander under a ± 5 percent variation of the modal frequencies. It is seen that two of the pole pairs are particularly sensitive and are nearly driven unstable by only this ± 5 percent variation. This happens because the associated structural modes contribute little to performance and the LQG design attempts a "cheap control" (small regulator and observer gains) by placing compensator poles very close to the open-loop plant poles. For nominal values, this scheme achieves significant shifts of open-loop poles with very small gains, but it is highly sensitive to off-nominal perturbations.

Figure 4.2-3 shows closed-loop poles for the same conditions except that a maximum entropy compensator design with $\sigma = 0.1$ (10 percent variation modelled) was utilized. In contrast with Figure 4.2-2, the maximum entropy design makes the compensator poles "stand-off" deeper in the left half-plane. (This is a direct consequence of the Stratonovich correction.) Consequently, the strong and sensitive interactions noted above are entirely eliminated. The poles associated with higher-order structural modes are seen to vary only along the imaginary axis and are not destabilized.

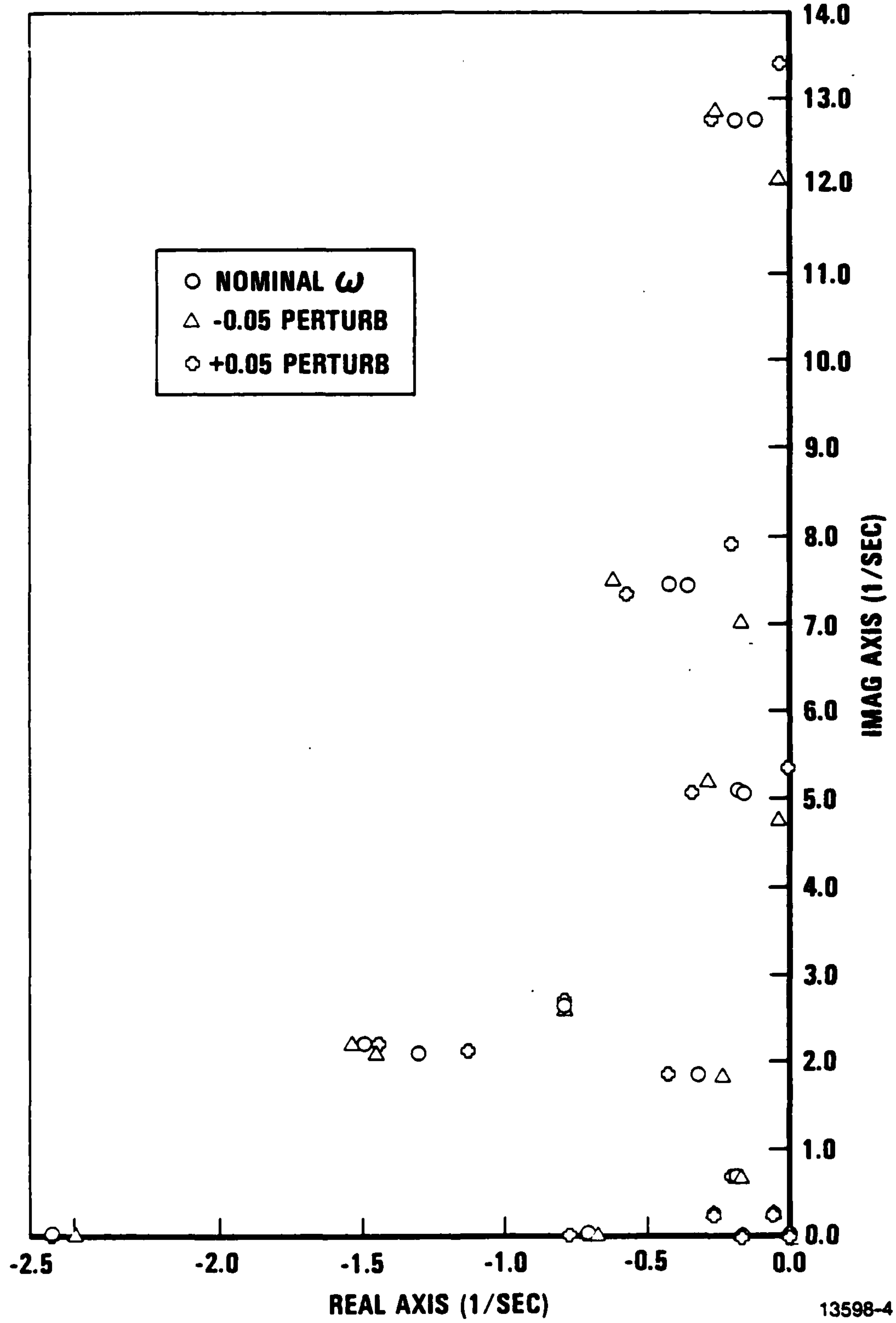


Figure 4.2-2. LQG Poles

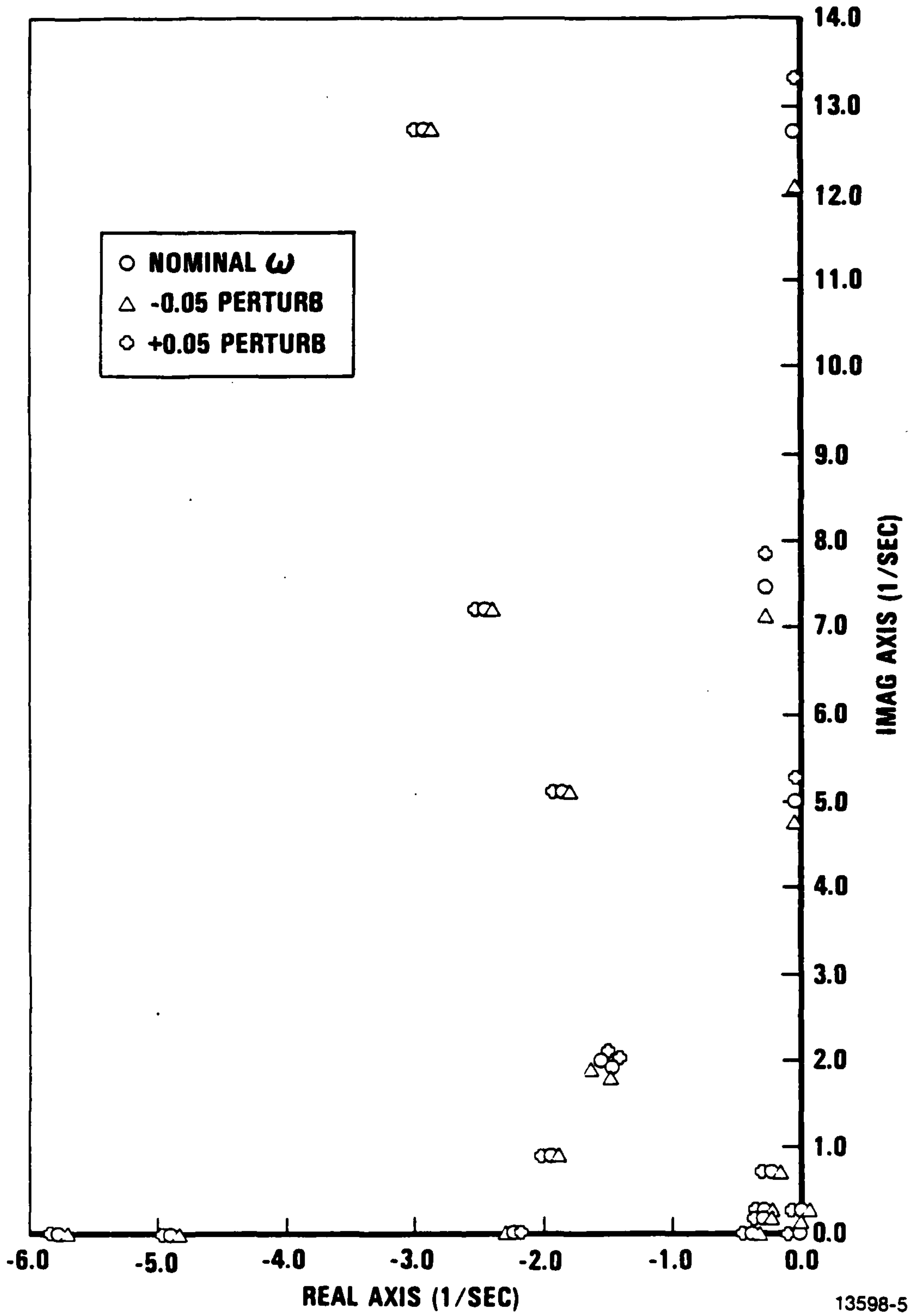


Figure 4.2-3. SIGMA = 0.1 Poles

Figure 4.2-4 illustrates how the total performance index for given controller designs varies as the structural mode frequencies are perturbed relative to their nominal values. The LQG design (which is simply a maximum entropy design for $\sigma = 0$) becomes unstable for > 7 percent and < -14 percent variations. In contrast and even with a modest 10 percent level of modelled uncertainty, the maximum entropy designs completely eliminate the sensitivity. Note that within the parameter range for which LQG is stable, the $\sigma = 0.1$ maximum entropy design experiences only 12-15 percent degradation. Of course, over the regions for which LQG is unstable, the maximum entropy designs are qualitatively superior.

These results serve to illustrate a general fact: By incorporating parameter uncertainty as an intrinsic facet of the basic design model, the maximum entropy formulation is able to secure high levels of robustness with little degradation of nominal performance.

Finally, the combined OP/ME design capability was exercised, taking the 16-state maximum entropy compensator design with $\sigma = 0.10$ frequency uncertainty level as the starting point. Reduced order compensator designs were constructed for compensators of order 14, 12, 10, 8, 6, and 5. Figure 4.2-5 shows the tradeoff between performance (total, closed-loop performance index evaluated for nominal values of modal frequencies) and controller dimension. The figure clearly shows that performance degradation for compensator orders above 6 is negligible. The 6th order controller sacrifices only 3 percent of the performance of the full-order (16-state) controller. This would seem to be acceptable in view of the better than sixfold decrease in implementation costs (e.g., flops required in matrix multiplication) which results from order reduction.

In conclusion, these results, together with much additional material included in [27], demonstrate automated solution of the full OP/ME design equations (shown in Figure 4.1-4) and illustrate the performance and implementation benefits to be expected under this unified approach.

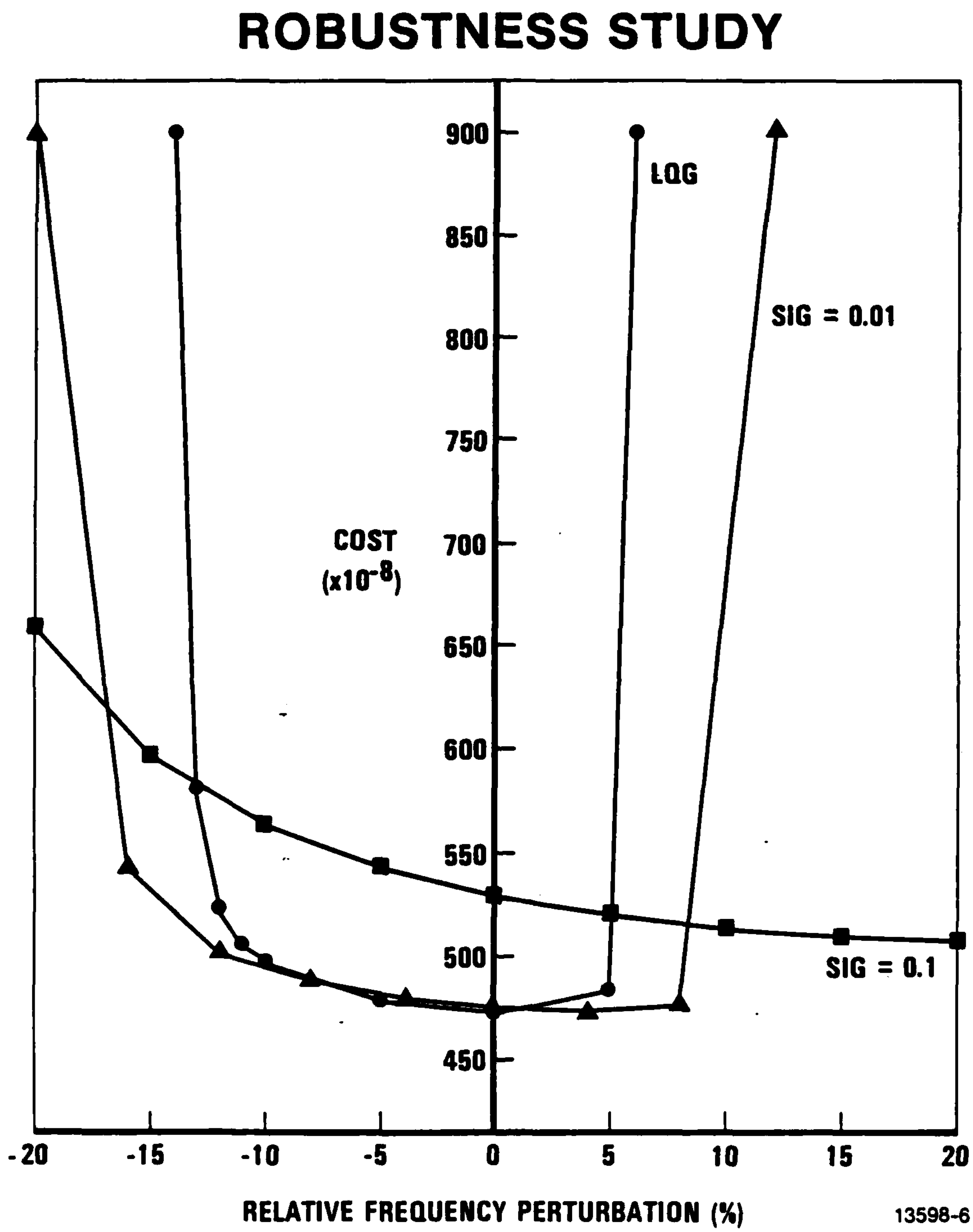


Figure 4.2-4. Robustness Study for SCOLE

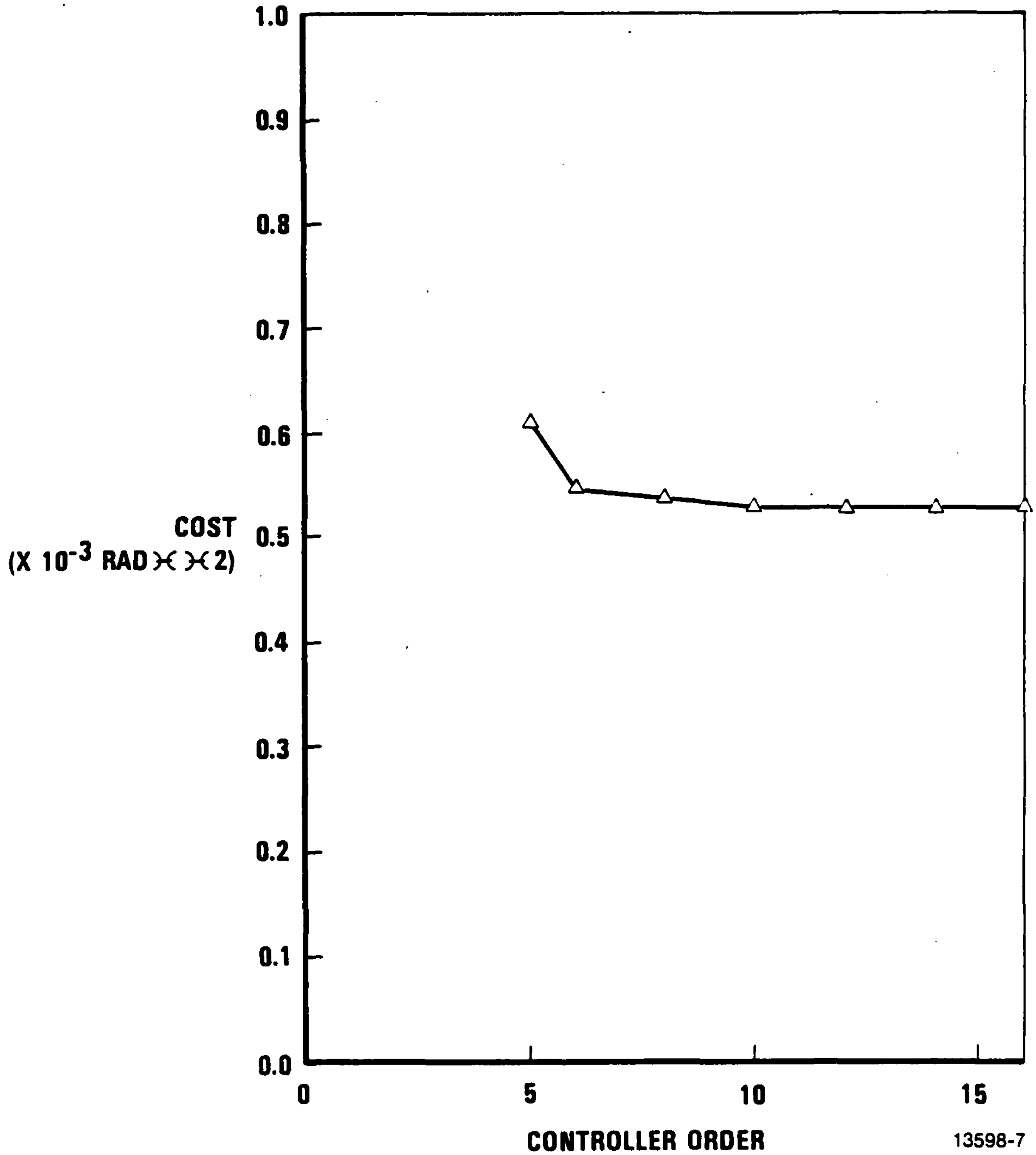


Figure 4.2-5. Cost Versus Complexity

References

Given herein is a comprehensive list containing all reference works utilized in the present study. These references are generally grouped by topic as follows:

Reference	Topic
[1] - [11]	Statistical energy analysis and early works on the maximum entropy modelling approach
[12] - [27]	The optimal projection formulation and the combined optimal projection/maximum entropy design approach; [18], [19], [21] and [27] present related spacecraft control applications studies
[28] - [30]	Extensions of the optimal projection formulation to distributed parameter systems and to reduced-order modelling and state estimation
[31] - [34]	Recent extensions of the combined optimal projection/maximum entropy design approach
[35] - [62]	General references on model reduction for linear systems
[63] - [80]	General references on the problem of reduced-order state estimator design
[81] - [110]	Design of optimal contrained-structure (e.g., reduced-order) dynamic compensators
[111] - [129]	Stabilizability and pole assignability for fixed-form dynamic compensation
[130] - [156]	General review of control theory for distributed parameter systems ([130], [131], [134] and [152]) and results on full- and fixed-order compensation and full-state feedback with related stabilizability results
[157] - [166]	General references on linear algebra, functional analysis and semigroup theory
[167] - [172]	Balancing transformations and generalized inverses with applications to distributed parameter systems
[173] - [182]	Optimal stochastic control theory and the maximum entropy formalism
[183] - [204]	General references on stochastic processes and stochastic differential equations
[205] - [236]	Optimal filtering and control theory for stochastic dynamic systems, including linear systems with stochastic multiplicative coefficients

[237] - [265] Stability and stabilizability results for stochastic systems

[266] - [273] Matrix theory and linear algebra with applications to optimization theory

[274] - [278] Miscellaneous references on systems with stochastic parameters and guaranteed cost control methodologies

[279] - [286] Stochastic systems theory for distributed parameter systems

[287] - [291] Miscellaneous works on optimization and control design for uncertain systems

[292] - [303] L^{∞}/H^{∞} optimization methods and robustness properties

[304] - [321] Stability robustness and sensitivity reduction for uncertain dynamical systems with bounded parameter variations

[322] - [323] Basic references on information theory

1. R. H. Lyon, Statistical Energy Analysis of Dynamical Systems: Theory and Applications, MIT Press, Cambridge, MA, 1975.

2. D. C. Hyland, "Optimal Regulation of Structural Systems With Uncertain Parameters," MIT Lincoln Laboratory, TR-551, 2 February 1981, DDC# AD-A099111/7.

3. D. C. Hyland, "Active Control of Large Flexible Spacecraft: A New Design Approach Based on Minimum Information Modelling of Parameter Uncertainties," Proc. Third VPI&SU/AIAA Symposium, pp. 631-646, Blacksburg, VA, June 1981.

4. D. C. Hyland, "Optimal Regulator Design Using Minimum Information Modelling of Parameter Uncertainties: Ramifications of the New Design Approach," Proc. Third VPI&SU/AIAA Symposium, pp. 701-716, Blacksburg, VA, June 1981.

5. D. C. Hyland and A. N. Madiwale, "Minimum Information Approach to Regulator Design: Numerical Methods and Illustrative Results," Proc. Third VPI&SU/AIAA Symposium, pp. 101-118, Blacksburg, VA, June 1981.

6. D. C. Hyland and A. N. Madiwale, "A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters," Proc. AIAA Guid. Contr. Conf., pp. 324-332, Albuquerque, NM, August 1981.

7. D. C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters," AIAA 20th Aerospace Sciences Meeting, paper 82-0312, Orlando, FL, January 1982.

8. D. C. Hyland, "Structural Modeling and Control Design Under Incomplete Parameter Information: The Maximum Entropy Approach," AFOSR/NASA Workshop in Modeling, Analysis and Optimization Issues for Large Space Structures, Williamsburg, VA, May 1982.

9. D. C. Hyland, "Minimum Information Stochastic Modelling of Linear Systems with a Class of Parameter Uncertainties," Proc. Amer. Contr. Conf., pp. 620-627, Arlington, VA, June 1982.

10. D. C. Hyland, "Maximum Entropy Stochastic Approach to Control Design for Uncertain Structural Systems," Proc. Amer. Contr. Conf., pp. 680-688, Arlington, VA, June 1982.

11. D. C. Hyland, "Minimum Information Modeling of Structural Systems with Uncertain Parameters," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 71-88, JPL, Pasadena, CA, July 1982.

12. D. C. Hyland and A. N. Madiwale, "Fixed-Order Dynamic Compensation Through Optimal Projection," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 409-427, JPL, Pasadena, CA, July 1982.

13. D. C. Hyland, "Mean-Square Optimal Fixed-Order Compensation - Beyond Spillover Suppression," paper 1403, AIAA Astrodynamics Conference, San Diego, CA, August 1982.

14. D. C. Hyland, "Robust Spacecraft Control Design in the Presence of Sensor/Actuator Placement Errors," AIAA Astrodynamics Conference, San Diego, CA, August 1982.

15. D. C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results," AIAA 21st Aerospace Sciences Meeting, paper 83-0303, Reno, NV, January 1983.

16. D. C. Hyland, "Mean-Square Optimal, Full-Order Compensation of Structural Systems with Uncertain Parameters," MIT, Lincoln Laboratory TR-626, 1 June 1983.

17. D. C. Hyland and D. S. Bernstein, "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation," Proc. 22nd IEEE Conf. Dec. Contr., pp. 161-165, San Antonio, TX, December 1983.

18. F. M. Ham, J. W. Shipley and D. C. Hyland, "Design of a Large Space Structure Vibration Control Experiment," Proc. 2nd Int. Modal Anal. Conf., pp. 550-558, Orlando, FL, February 1984.

19. D. C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," Proc. AIAA Dynamics Specialists Conf., pp. 381-389, Palm Springs, CA, May 1984.

20. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation of Distributed Parameter Systems," Proc. AIAA Dynamics Specialists Conf., pp. 396-400, Palm Springs, CA, May 1984.

21. F. M. Ham and D. C. Hyland, "Vibration Control Experiment Design for the 15-M Hoop/Column Antenna," Proceedings of the Workshop on Identification and Control of Flexible Space Structures, San Diego, CA, June 1984.

22. D. S. Bernstein and D. C. Hyland, "Numerical Solution of the Optimal Model Reduction Equations," Proc. AIAA Guid. Contr. Conf., pp. 560-562, Seattle, WA, August 1984.

23. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 1034-1037, 1984.

24. D. C. Hyland, "Application of the Maximum Entropy/Optimal Projection Control Design Approach for Large Space Structures," Large Space Antenna Systems Technology Conference, NASA Langley, December 1984.

25. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore," 23rd IEEE Conf. Dec. Contr., pp. 120-126, Las Vegas, NV, December 1984.

26. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Approach to Designing Optimal Finite-Dimensional Controllers for Distributed Parameter Systems," Proc. 23rd IEEE Conf. Dec. Contr., pp. 556-560, Las Vegas, NV, December 1984 (also presented at the SIAM Fall Meeting, Norfolk, VA, November 1983).

27. L. D. Davis, D. C. Hyland and D. S. Bernstein, "Application of the Maximum Entropy Design Approach to the Spacecraft Control Laboratory Experiment (SCOLE)," Final Report, NASA Langley, January 1985.

28. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," SIAM J. Contr. Optim., to appear.

29. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," IEEE Trans. Autom. Contr., to appear.

30. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," submitted for publication.

31. D. S. Bernstein and D. C. Hyland, "Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis" (abstract), submitted for publication.

32. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order Modelling, State Estimation and Dynamic Compensation of Linear Systems with Multiplicative Stratonovich White Noise," in preparation.

33. D. S. Bernstein, L. D. Davis, F. M. Ham and D. C. Hyland, "The Optimal Projection Equations for Fixed-Order Modelling, Estimation and Control of Discrete-Time Stochastic Bilinear Systems," in preparation.

34. D. C. Hyland, "Guaranteed Performance and Stability Robustness via Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis," in preparation.

35. E. J. Davison, "A Method for Simplifying Linear Dynamic Systems," IEEE Trans. Autom. Contr., Vol. AC-11, pp. 93-101, 1966.

36. M. Aoki, "Control of Large-Scale Dynamic Systems by Aggregation," IEEE Trans. Autom. Cont., Vol. AC-13, pp. 246-253, 1968.

37. D. Mitra, "Analytical Results on the Use of Reduced Models in the Control of Linear Dynamical Systems," Proc. IEE, Vol. 116, pp. 1439-1444, 1969.

38. D. A. Wilson, "Optimum Solution of Model-Reduction Problem," Proc. IEE, Vol. 117, pp. 1161-1165, 1970.

39. J. D. Aplevich, "Approximation of Discrete Linear Systems," Int. J. Contr., pp. 565-575, 1973.

40. D. A. Wilson, "Model Reduction for Multivariable Systems," Int. J. Contr., Vol. 20, pp. 57-64, 1974.

41. M. F. Hutton and B. Friedland, "Routh Approximations for Reducing the Order of Linear, Time-Invariant Systems," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 329-337, 1975.

42. R. Genesio and M. Milanese, "A Note on the Derivation and Use of Reduced-Order Models," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 118-122, 1976.

43. J. Hickin and N. K. Sinha, "Applications of Projective Reduction Methods to Estimation and Control," J. Cyber., Vol. 8, pp. 159-181, 1978.

44. E. C. Y. Tse, J. V. Medanic and W. R. Perkins, "Generalized Hessenberg Transformation for Reduced-Order Modelling of Large-Scale Systems," Int. J. Contr., Vol. 27, pp. 493-512, 1978.

45. A. Arbel and E. Tse, "Reduced-Order Models, Canonical Forms and Observers," Int. J. Contr., Vol. 30, pp. 513-531, 1979.

46. R. N. Mishra and D. A. Wilson, "A New Algorithm for Optimal Reduction of Multivariable Systems," Int. J. Contr., Vol. 31, pp. 443-466, 1980.

47. P. T. Kabamba, "Model Reduction by Euclidean Methods," J. Guid. Contr., Vol. 3, pp. 555-562, 1980.

48. Y. Baram and Y. Be'eri, "Stochastic Model Simplification," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 379-390, 1981.

49. S.-Y. Kung and D. W. Lin, "Optimal Hankel-Norm Model Reductions: Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 832-852, 1981.

50. B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 17-32, 1981.

51. E. Eitelberg, "Model Reduction by Minimizing the Weighted Equation Error," Int. J. Contr., Vol. 34, pp. 1113-1123, 1981.

52. R. E. Skelton and A. Yousuff, "Component Cost Analysis of Large Scale Systems," in Control and Dynamic Systems, Vol. 18, C. T. Leondes, ed., Academic Press, 1982.

53. D. Bonvin and D. A. Mellichamp, "A Unified Derivation and Critical Review of Modal Approaches to Model Reduction," Int. J. Contr., Vol. 35, pp. 829-848, 1982.

54. J. K. Tugnait, "Continuous-Time Stochastic Model Simplification," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 993-996, 1982.

55. C. S. Sims, "Reduced-Order Modelling and Filtering," in Control and Dynamic Systems, Vol. 18, pp. 55-103, C. T. Leondes, ed., 1982.

56. L. Pernebo and L. M. Silverman, "Model Reduction via Balanced State Space Representations," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 382-387, 1982.

57. K. V. Fernando and H. Nicholson, "On the Structure of Balanced and Other Principal Representations of SISO Systems," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 228-231, 1983.

58. S. Shokoohi, L. M. Silverman and P. M. Van Dooren, "Linear Time-Variable Systems: Balancing and Model Reduction," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 810-822, 1983.

59. E. I. Verriest and T. Kailath, "On Generalized Balanced Realizations," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 833-844, 1983.

60. R. E. Skelton and A. Yousuff, "Component Cost Analysis of Large Scale Systems," Int. J. Contr., Vol. 37, pp. 285-304, 1983.

61. L. J. Shieh and Y. T. Tsay, "Algebra-Geometric Approach for the Model Reduction of Large-Scale Multivariable Systems," Proc. IEE, Vol. 131, pp. 23-36, 1984.

62. P. T. Kabamba, "Balanced Gains and Their Significance for Balanced Model Reduction," Proc. Conf. Inform. Sci. Sys., Princeton Univ., 1984.

63. D. G. Luenberger, "Observers for Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-11, pp. 190-197, 1966.

64. K. W. Simon and A. R. Stubburud, "Reduced Order Kalman Filter," Int. J. Contr., Vol. 10, pp. 501-509, 1969.

65. C. S. Sims and J. L. Melsa, "Specific Optimal Estimation," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 183-186, 1969.

66. C. S. Sims and J. L. Melsa, "A Survey of Specific Optimal Techniques in Control and Estimation," Int. J. Contr., pp. 299-308, 1971.

67. T. Yoshikawa, "Minimal-Order Optimal Filters for Discrete-Time Linear Stochastic Systems," Int. J. Contr., Vol. 21, pp. 1-19, 1975.

68. R. B. Asher, K. D. Herring and J. C. Ryles, "Bias Variance and Estimation Error in Reduced Order Filters," Automatica, Vol. 12, pp. 589-600, 1976.

69. J. I. Galdos and D. E. Gustafson, "Information and Distortion in Reduced-Order Filter Design," IEEE Trans. Inform. Thy., Vol. IT-23, pp. 183-194, 1977.

70. F. W. Fairman, "Reduced Order State Estimators for Discrete-Time Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 673-675, 1977.

71. F. W. Fairman, "On Stochastic Observer Estimators for Continuous-Time Systems," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 874-876, 1977.

72. C. S. Sims, "Optimal and Suboptimal Results in Full- and Reduced-Order Linear Filtering," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 469-472, 1978.

73. C. S. Sims and L. G. Stotts, "Linear Discrete Reduced-Order Filtering," Proc. IEEE Conf. Dec. Contr., 1979.

74. D. A. Wilson and R. N. Mishra, "Design of Low Order Estimators Using Reduced Models," Int. J. Contr., Vol. 23, pp. 447-456, 1979.

75. K. Yonezawa, "Reduced-Order Kalman Filtering With Incomplete Observability," J. Guid. Contr., Vol. 3, pp. 270-282, 1980.

76. F. W. Fairman and R. D. Gupta, "Design of Multifunctional Reduced-Order Observers," Int. J. Sys. Sci., Vol. 11, pp. 1083-1094, 1980.

77. E. Fogel and Y. F. Huang, "Reduced-Order Optimal State Estimator for Linear Systems with Partially Noise Corrupted Measurements," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 994-996, 1980.

78. M. H. Dwarakanath, "A Proof of the Minimal Order Observer," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 998-1000, 1982.

79. U. V. Dombrovskii, "Method of Synthesizing Suboptimal Filters of Reduced Order for Digital Linear Dynamic Systems," Autom. Remote Contr., Vol. 43, pp. 1483-1489, 1982.

80. T. Hinamoto and F. W. Fairman, "Reduced Order Observer Design for a Linear Map of the State," J. Franklin Inst., Vol. 314, pp. 95-108, 1982.

81. T. L. Johnson and M. Athans, "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 658-660, 1970.

82. W. S. Levine, T. L. Johnson and M. Athans, "Optimal Limited State Variable Feedback Controllers for Linear Systems," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 785-793, 1971.

83. K. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, New York, 1972.

84. D. B. Rom and P. E. Sarachik, "The Design of Optimal Compensators for Linear Constant Systems with Inaccessible States," IEEE Trans. Autom. Contr., Vol. AC-18, pp. 509-512, 1973.

85. M. Sidar and B. Z. Kurtaran, "Optimal Low-Order Controllers for Linear Stochastic Systems," Int. J. Contr., Vol. 22, 377-387, 1975.

86. J. M. Mendel and J. Feather, "On the Design of Optimal Time-Invariant Compensators for Linear Stochastic Time-Invariant Systems," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 653-657, 1975.

87. S. Basuthakur and C. H. Knapp, "Optimal Constant Controllers for Stochastic Linear Systems," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 664-666, 1975.

88. R. B. Asher and J. C. Durrett, "Linear Discrete Stochastic Control with a Reduced-Order Dynamic Compensator," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 626-627, 1976.

89. G. N. Mil'shtein, "Linear Optimal Controllers of Specified Structures in Systems with Incomplete Information," Autom. Remote Contr., Vol. 37, pp. 1179-1183, 1976.

90. W. J. Naeije and O. H. Bosgra, "The Design of Dynamic Compensators for Linear Multivariable Systems," 1977 IFAC, Fredricton, N.B., Canada, pp. 205-212.

91. H. R. Sirisena and S. S. Choi, "Design of Optimal Constrained Dynamic Compensators for Non-Stationary Linear Stochastic Systems," Int. J. Contr., Vol. 25, pp. 513-524, 1977.

92. P. J. Blanvillain and T. L. Johnson, "Invariants of Optimal Minimal-Order Observer-Based Compensators," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 473-474, 1978.

93. P. J. Blanvillain and T. L. Johnson, "Specific-Optimal Control With a Dual Minimal-Order Observer-Based Compensator," Int. J. Contr., Vol. 28, pp. 277-294, 1978.

94. C. J. Wenk and C. H. Knapp, "Parameter Optimization in Linear Systems with Arbitrarily Constrained Controller Structure," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 496-500, 1980.

95. J. O'Reilly, "Optimal Low-Order Feedback Controllers for Linear Discrete-Time Systems," in Control and Dynamic Systems, Vol. 16, C. T. Leondes, ed., Academic Press, 1980.

96. D. P. Looze and N. R. Sandell, Jr., "Gradient Calculations for Linear Quadratic Fixed Control Structure Problems," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 258-285, 1980.

97. M. J. Grimble, "Reduced-Order Optimal Controller for Discrete-Time Stochastic Systems," Proc. IEE, Vol. 127, pp. 55-63, 1980.

98. P. T. Kabamba and R. W. Longman, "An Integrated Approach to Reduced-Order Control Theory," Optim. Contr. Appl. Meth., Vol. 4, pp. 405-415, 1983.

99. E. A. Jonckheere and L. M. Silverman, "A New Set of Invariants for Linear Systems - Application to Reduced-Order Compensator Design," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 953-964, 1983.

100. A. Yousuff and R. E. Skelton, "A Note on Balanced Controller Reduction," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 254-257, 1984.

101. A. Yousuff and R. E. Skelton, "Controller Reduction by Component Cost Analysis," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 520-530, 1984.

102. A. Yousuff and R. E. Skelton, "Covariance Equivalent Realizations with Application to Model Reduction of Large Scale Systems," in Control and Dynamic Systems, Vol. 22, C. T. Leondes, ed., 1984.

103. W. S. Levine and M. Athans, "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 44-48, 1970.

104. A. Jameson, "Optimization of Linear Systems of Constrained Configuration," Int. J. Contr., Vol. 11, pp. 409-421, 1970.

105. R. L. Kosut, "Suboptimal Control of Linear Time-Invariant Systems Subject to Control Structure Constraints," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 557-563, 1970.

106. C. H. Knapp and S. Basuthakur, "On Optimal Output Feedback," IEEE Trans. Autom. Contr., Vol. AC-17, pp. 823-825, 1972.

107. J. M. Mendel, "A Concise Derivation of Optimal Constant Limited State Feedback Gains," IEEE Trans. Autom. Contr., Vol. 19, pp. 447-448, 1974.

108. S. P. Bingulac, N. M. Cuk and M. S. Calovic, "Calculation of Optimum Feedback Gains for Output-Constrained Regulators," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 164-166, 1975.

109. J. C. Allwright, "Optimal Output Feedback Without Trace," Appl. Math. Optim., Vol. 2, pp. 351-372, 1976.

110. Dj. B. Petkovski and M. Rakic, "On the Calculation of Optimum Feedback Gains for Output-Constrained Regulators," IEEE Trans. Autom. Contr., Vol. AC-23, p. 260, 1978.

111. J. Medanic, "On Stabilization and Optimization by Output Feedback," Proc. Twelfth Asilomar Conf. Circ., Sys. Comp., pp. 412-416, 1978.

112. W. E. Hopkins, Jr., J. Medanic and W. R. Perkins, "Output Feedback Pole Placement in the Design of Suboptimal Linear Quadratic Regulators," Int. J. Contr., Vol. 34, pp. 593-612, 1981.

113. S. Renjen and D. P. Looze, "Synthesis of Decentralized Output/State Regulators," Proc. Amer. Contr. Conf., pp. 758-762, 1982.

114. J. B. Pearson, "Compensator Design for Dynamic Compensation," Int. J. Contr., Vol. 9, pp. 473-482, 1969.

115. J. B. Pearson and C. Y. Ding, "Compensator Design for Multivariable Linear Systems," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 130-134, 1969.

116. F. M. Brasch and J. B. Pearson, "Pole Placement Using Dynamic Compensators," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 34-43, 1970.

117. R. Ahmari and A. C. Vacroux, "On the Pole Assignment in Linear Systems with Fixed-Order Compensators," Int. J. Contr., Vol. 17, pp. 397-404, 1973.

118. D. C. Youla, J. J. Bongiorno, Jr. and C. N. Lu, "Single-Loop Feedback-Stabilization of Linear Multivariable Dynamical Plants," Automatica, Vol. 10, pp. 159-173, 1974.

119. H. Seraji, "An Approach to Dynamic Compensator Design for Pole Assignment," Int. J. Contr., Vol. 21, pp. 955-966, 1975.

120. R. V. Patel, "Design of Dynamic Compensators for Pole Assignment," Int. J. Sys. Sci., Vol. 7, pp. 207-224, 1976.

121. D. C. Youla, J. J. Bongiorno, Jr. and H. A. Jabr, "Modern Wiener-Hopf Design of Optimal Controllers Part I: The Single-Input-Output Case," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 3-13, 1976.

122. D. C. Youla, H. A. Jabr and J. J. Bongiorno, Jr., "Modern Wiener-Hopf Design of Optimal Controllers Part II: The Multivariable Case," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 319-338, 1976.

123. L. Chang and J. B. Pearson, "Frequency Domain Synthesis of Multivariable Linear Regulators," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 3-15, 1978.

124. C. A. Desoer, R. W. Liu, J. Murray and R. Saeks, "Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 399-412, 1980.

125. H. Seraji, "Design of Pole-Placement Compensators for Multivariable Systems," Automatica, Vol. 16, pp. 335-338, 1980.

126. H. Imai and H. Akashi, "Disturbance Localization and Pole Shifting by Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 226-235, 1981.

127. H. Seraji, "On Fixed Modes in Decentralized Control Systems," Int. J. Contr., Vol. 35, pp. 775-784, 1982.

128. M. Vidyasagar, H. Schneider and B. A. Francis, "Algebraic and Topological Aspects of Feedback Stabilization," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 880-894, 1982.

129. M. Vidyasagar and N. Viswanadham, "Algebraic Design Techniques for Reliable Stabilization," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 1085-1095, 1982.

130. M. Athans, "Toward a Practical Theory of Distributed Parameter Systems," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 245-247, 1970.

131. M. J. Balas, "Toward a More Practical Control Theory for Distributed Parameter Systems," in Control and Dynamic Systems, Vol. 19, pp. 361-418, C. T. Leondes, ed., Academic Press, New York, 1982.

132. R. F. Curtain and A. J. Pritchard, Infinite Dimensional Linear Systems Theory, Springer-Verlag, New York, 1978.

133. J. S. Gibson, "The Riccati Integral Equations for Optimal Control Problems on Hilbert Spaces," SIAM J. Contr. Optim., Vol. 17, pp. 537-565, 1979.

134. A. V. Balakrishnan, Applied Functional Analysis, Springer-Verlag, New York, 1981.

135. J. S. Gibson, "An Analysis of Optimal Modal Regulation: Convergence and Stability," SIAM J. Contr. Optim., Vol. 19, pp. 686-707, 1981.

136. J. S. Gibson, "Linear-Quadratic Optimal Control of Hereditary Differential Systems: Infinite Dimensional Riccati Equations and Numerical Approximations," SIAM J. Contr. Optim., Vol. 21, pp. 95-139, 1983.

137. H. T. Banks and K. Kunisch, "The Linear Regulator Problem for Parabolic Systems," SIAM J. Contr. Optim., Vol. 22, pp. 684-698, 1984.

138. H. T. Banks, K. Ito and I. G. Rosen, "A Spline Based Technique for Computing Riccati Operators and Feedback Controls in Regulator Problems for Delay Equations," SIAM J. Sci. Stat. Comput., Vol. 5, 1984.

139. T. L. Johnson, "Optimization of Low Order Compensators for Infinite Dimensional Systems," Proc. 9th IFIP Symp. Optim. Tech., pp. 394-401, Warsaw, Poland, September 1979.

140. R. K. Pearson, "Optimal Fixed-Form Compensators for Large Space Structures," in ACOSS SIX (Active Control of Space Structures), RADC-TR-81-289, Final Technical Report, Rome Air Development Center, Griffiss AFB, New York, 1981.

141. R. K. Pearson, Optimal Velocity Feedback Control of Flexible Structures, Ph.D. Dis., MIT Dept. Elec. Eng. Comp. Sci., 1982.

142. R. F. Curtain, "Compensators for Infinite-Dimensional Linear Systems," J. Franklin Inst., Vol. 315, pp. 331-346, 1983.

143. J. M. Schumacher, "A Direct Approach to Compensator Design for Distributed Parameter Systems," SIAM J. Contr. Optim., Vol. 21, pp. 823-836, 1983.

144. C. N. Nett, C. A. Jacobson and M. J. Balas, "Fractional Representation Theory: Robustness Results With Applications to Finite Dimensional Control of a Class of Linear Distributed Systems," Proc. IEEE Conf. Dec. Contr., pp. 268-280, San Antonio, TX, December 1983.

145. M. J. Balas, "Linear Distributed Parameter Systems: Closed-Loop Exponential Stability with a Finite-Dimensional Controller," Automatica, Vol. 20, pp. 371-377, 1984.

146. C. Jacobson, "Necessary and Sufficient Conditions for Output Feedback Stabilization of a Class of Linear Distributed Systems," preprint.

147. D. L. Russel, "Linear Stabilization of the Linear Oscillator in Hilbert Space," J. Math. Anal. Appl., Vol. 25, pp. 663-675, 1969.

148. D. L. Russell, "Decay Rates for Weakly Damped Systems in Hilbert Space Obtained With Control Theoretic Methods," J. Diff. Eqns., Vol. 19, pp. 344-370, 1975.

149. M. J. Balas, "Modal Control of Certain Flexible Dynamic Systems," SIAM J. Contr. Optim., Vol. 16, pp. 450-462, 1978.

150. M. J. Balas, "Feedback Control of Flexible Systems," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 673-679, 1978.

151. J. S. Gibson, "A Note on Stabilization of Infinite Dimensional Linear Oscillators by Compact Feedback," SIAM J. Contr. Optim., Vol. 18, pp. 311-316, 1980.

152. M. J. Balas, "Trends in Large Space Structure Control Theory: Fondest Hopes, Wildest Dreams," IEEE Trans. Autom. Contr., Vol. AC-24, pp. 522-535, 1982.

153. M. J. Balas, "Reduced-Order Feedback Control of Distributed Parameter Systems via Singular Perturbation Methods," J. Math. Anal. Appl., Vol. 87, pp. 281-294, 1982.

154. M. J. Balas, "The Galerkin Method and Feedback Control of Linear Distributed Parameter Systems," J. Math. Anal. Appl., Vol. 91, pp. 529-546, 1983.

155. T. L. Johnson, "Progress in Modelling and Control of Flexible Spacecraft," J. Franklin Inst., Vol. 315, pp. 495-520, 1983.

156. M. J. Balas, "Feedback Control of Dissipative Hyperbolic Distributed Parameter Systems with Finite Dimensional Controllers," J. Math. Anal. Appl., Vol. 98, pp. 1-24, 1984.

157. J. R. Ringrose, Compact Non-Self-Adjoint Operators, Van Nostrand Reinhold Co., London, 1971.

158. I. C. Gohberg and M. G. Krein, Introduction to the Theory of Linear Non-Self-Adjoint Operators, Translations of Mathematical Monographs, Vol. 18, American Mathematical Society, Providence, RI, 1966.

159. I. Gohberg and S. Goldberg, Basic Operator Theory, Birkhauser, Boston, 1981.

160. F. R. Gantmacher, The Theory of Matrices, Vol. I, Chelsea, NY, 1977.

161. B. Noble and J. W. Daniel, Applied Linear Algebra, Second Edition, Prentice-Hall, Englewood Cliffs, NJ, 1977.

162. E. Hille and R. S. Phillips, Functional Analysis and Semigroups, Colloq. Publ. Vol. 31, Amer. Math. Soc., Providence, RI, 1957.

163. T. Kato, Perturbation Theory for Linear Operations, Second Edition, Springer-Verlag, New York, 1976.

164. R. F. Curtain and A. J. Pritchard, Functional Analysis in Modern Applied Mathematics, Academic Press, London, 1977.

165. N. U. Ahmed and K. L. Teo, Optimal Control of Distributed Parameter Systems, North-Holland, New York, 1981.

166. A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, 1983.

167. S. Chakrabarti, B. B. Battacharyya and M. N. S. Swamy, "On Simultaneous Diagonalization of a Collection of Hermitian Matrices," The Matrix and Tensor Quarterly, Vol. 29, pp. 35-54, 1978.

168. A. J. Laub, "Computation of Balancing Transformation," Proc. 1980 Joint Autom. Contr. Conf., San Francisco, CA, Aug. 1980.

169. D. C. Lay, "Spectral Properties of Generalized Inverses of Linear Operators," SIAM J. Appl. Math., Vol. 29, pp. 103-109, 1975.

170. P. Robert, "On the Group-Inverse of a Linear Transformation," J. Math. Anal. Appl., Vol. 22, pp. 658-669, 1968.

171. R. F. Curtain, "Finite-Dimensional Compensators for Parabolic Distributed Systems With Unbounded Control and Observation," SIAM J. Contr. Optim., Vol. 22, pp. 255-276, 1984.

172. M. J. Balas, "The Structure of Discrete-Time Finite-Dimensional Control of Distributed Parameter Systems," Proc. IEEE Int. Large Scale Sys. Symp., Virginia Beach, VA, 1982.

173. Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. Autom. Contr., Vol. AC-19, pp. 494-500, 1974.

174. Y. Bar-Shalom and E. Tse, "Generalized Certainty Equivalence and Dual Effect in Stochastic Control," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 817-819, 1975.

175. Y. Bar-Shalom and E. Tse, "Caution, Probing and the Value of Information in the Control of Uncertain Systems," Annals of Economic and Social Measurement, Vol. 5, pp. 323-337, 1976.

176. E. T. Jaynes, "Information Theory and Statistical Mechanics," Phys. Rev., Vol. 106, pp. 620-630, 1957.

177. E. T. Jaynes, "New Engineering Applications of Information Theory," Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability, J. L. Bogdanoff and F. Kozin, pp. 163-203, Wiley, New York, 1963.

178. E. T. Jaynes, "Prior Probabilities," IEEE Trans. Sys. Sci. Cybern., Vol. SSC-4, pp. 227-241, 1968.

179. E. T. Jaynes, "Where Do We Stand on Maximum Entropy?," in The Maximum Entropy Formalism, D. Levine and M. Tribus, eds., MIT Press, pp. 15-118, Cambridge, MA, 1979.

180. R. D. Rosenkrandtz, ed., E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics, Reidel, Boston, 1983.

181. R. Christensen, Entropy Minimax Sourcebook, Vol. 1, Entropy Limited, Lincoln, MA, 1981.

182. E. Bonomi and J.-L. Lutton, "The N-City Travelling Salesman Problem: Statistical Mechanics and the Metropolis Algorithm," SIAM Review, Vol. 26, pp. 551-568, 1984.

183. K. Ito, On Stochastic Differential Equations, Amer. Math. Soc., Providence, RI, 1951.

184. I. M. Gelfand and A. M. Yaglom, "Calculation of the Amount of Information About a Random Function Contained in Another Such Function," Trans. Amer. Math. Soc., Vol. 18, pp. 199-245, 1959.

185. M. S. Pinsker, Information and Information Stability of Random Variables and Processes, Holden-Day, San Francisco, 1964.

186. E. Wong and M. Zakai, "On the Relation Between Ordinary and Stochastic Differential Equations," Int. J. Engrg. Sci., Vol. 3, pp. 213-229, 1965.

187. E. Wong and M. Zakai, "On the Convergence of Ordinary Integrals to Stochastic Integrals," Ann. Math. Statist., Vol. 36, pp. 1560-1564, 1965.

188. R. L. Stratonovich, "A New Representation for Stochastic Integrals," SIAM J. Contr., Vol. 4, pp. 362-371, 1966.

189. R. L. Stratonovich, Conditional Markov Process and Their Application to the Theory of Optimal Control, Elsevier, New York, 1968.

190. P. J. McLane, "Asymptotic Stability of Linear Autonomous Systems with State-Dependent Noise," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 754-755, 1969.

191. M. Zakai, "A Lyapunov Criterion for the Existence of Stationary Probability Distributions for Systems Perturbed by Noise," SIAM J. Contr., Vol. 7, pp. 390-397, 1969.

192. A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

193. E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

194. E. J. McShane, Stochastic Calculus and Stochastic Models, Academic Press, New York, 1974.

195. L. Arnold, Stochastic Differential Equations: Theory and Applications, Wiley, New York, 1974.

196. A. V. Balakrishnan, "On the Approximation of Ito Integrals Using Band-Limited Processes," SIAM. J. Contr. Optim., Vol. 12, pp. 237-251, 1974.

197. W. H. Fleming and R. W. Rishel, Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 1975.

198. E. J. McShane, "Stochastic Differential Equations," J. Multiv. Anal., Vol. 5, pp. 121-177, 1975.

199. H. J. Sussmann, "An Interpretation of Stochastic Differential Equations as Ordinary Differential Equations Which Depend on the Sample Point," Bull. Amer. Math. Soc., Vol. 83, pp. 296-298, 1977.

200. H. J. Sussmann, "On the Gap Between Deterministic and Stochastic Ordinary Differential Equations," Ann. Prob., Vol. 6, pp. 19-41, 1978.

201. M. Metivier and J. Pellaumail, Stochastic Integration, Academic Press, New York, 1980.

202. G. Kallianpur, Stochastic Filtering Theory, Springer-Verlag, New York, 1980.

203. N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.

204. A. Bagchi, "Approximation of Ito Integrals Arising in Stochastic Time-Delayed Systems," SIAM J. Contr. Optim., Vol. 22, pp. 878-888, 1984.

205. W. M. Wonham, "Lyapunov Criteria for Weak Stochastic Stability," J. Diff. Eqns., Vol. 2, pp. 195-207, 1966.

206. W. M. Wonham, "Optimal Stationary Control of Linear Systems With State-Dependent Noise," SIAM J. Contr., Vol. 5, pp. 486-500, 1967.

207. W. M. Wonham, "On a Matrix Riccati Equation of Stochastic Control," SIAM J. Contr., Vol. 6, pp. 681-697, 1968.

208. D. Kleinman, "Optimal Stationary Control of Linear Systems With Control-Dependent Noise," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 673-677, 1969.

209. P. J. McLane, "Optimal Linear Filtering for Linear Systems With State-Dependent Noise," Int. J. Contr., Vol. 10, pp. 41-51, 1969.

210. W. M. Wonham, "Random Differential Equations in Control Theory," in Probabilistic Analysis in Applied Mathematics, Vol. 2, pp. 131-212, A. T. Bharucha-Reid, ed., Academic Press, New York, 1970.

211. P. McLane, "Optimal Stochastic Control of Linear Systems with State- and Control-Dependent Disturbances," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 793-798, 1971.

212. D. Kleinman, "Numerical Solution of the State Dependent Noise Problem," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 419-420, 1976.

213. U. Haussmann, "Optimal Stationary Control With State and Control Dependent Noise," SIAM J. Contr., Vol. 9, pp. 184-198, 1971.

214. J. Bismut, "Linear-Quadratic Optimal Stochastic Control With Random Coefficients," SIAM J. Contr., Vol. 14, pp. 419-444, 1976.

215. A. Ichikawa, "Optimal Control of a Linear Stochastic Evolution Equation With State and Control Dependent Noise," Proc. IMA Conference on Recent Theoretical Developments in Control, pp. 383-401, Leicester, England, Academic Press, New York, 1976.

216. J. M. Bismut, "Linear Quadratic Optimal Stochastic Control With Random Coefficients," SIAM J. Contr. Optim., Vol. 14, pp. 419-444, 1976.

217. J. M. Bismut, "On Optimal Control of Linear Stochastic Equations With a Linear-Quadratic Criterion," SIAM J. Contr. Optim., Vol. 15, pp. 1-4, 1977.

218. U. G. Haussmann, "Asymptotic Stability of the Linear Ito Equation in Infinite Dimensions," J. Math. Anal. Appl., Vol. 65, pp. 219-235, 1978.

219. E. Pardoux, "Stochastic Partial Differential Equations and Filtering of Diffusion Processes," Stochastics, Vol. 3, pp. 127-167, 1979.

220. A. Ichikawa, "Dynamic Programming Approach to Stochastic Evolution Equations," SIAM J. Contr. Optim., Vol. 17, pp. 152-174, 1979.

221. H. V. Panossian and C. T. Leondes, "Observers for Optimal Estimation of the State of Linear Stochastic Discrete Systems," Int. J. Contr., Vol. 37, pp. 645-655, 1983.

222. H. Panossian and C. T. Leondes, "On Discrete-Time Riccati-Like Matrix Difference Equations with Random Coefficients," Int. J. Sys. Sci., Vol. 14, pp. 385-407, 1983.

223. N. U. Ahmed, "Stochastic Control on Hilbert Space for Linear Evolution Equations With Random Operator-Valued Coefficients," SIAM J. Contr. Optim., Vol. 19, pp. 401-430, 1981.

224. C. W. Merriam III, Automated Design of Control Systems, Gordon and Breach, New York, 1974.

225. M. Aoki, "Control of Linear Discrete-Time Stochastic Dynamic Systems With Multiplicative Disturbances," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 388-392, 1975.

226. D. E. Gustafson and J. L. Speyer, "Design of Linear Regulators for Nonlinear Systems," J. Spacecraft and Rockets, Vol. 12, pp. 351-358, 1975.

227. D. E. Gustafson and J. L. Speyer, "Linear Minimum Variance Filters Applied to Carrier Tracking," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 65-73, 1976.

228. Y. Sunahara and K. Yamashita, "An Approximate Method of State Estimation for Nonlinear Dynamical Systems With State-Dependent Noise," Int. J. Contr., Vol. 11, pp. 957-972, 1970.

229. C. S. Sims, "Discrete Reduced-Order Filtering With State-Dependent Noise," Proc. Joint Autom. Contr. Conf., 1980.

230. G. N. Mil'shtein, "Design of Stabilizing Controller with Incomplete State Data for Linear Stochastic System with Multiplicative Noise," Autom. Remote Contr., Vol. 43, pp. 653-659, 1982.

231. S. I. Marcus, "Modeling and Analysis of Stochastic Differential Equations Driven by Point Processes," IEEE Trans. Inform. Thy., Vol. IT-24, pp. 164-172, 1978.

232. H. J. Kushner, "Jump-Diffusion Approximations for Ordinary Differential Equations with Wide-Band Random Right Hand Sides," SIAM J. Contr. Optim., Vol. 17, pp. 729-744, 1979.

233. S. I. Marcus, "Modeling and Approximations of Stochastic Differential Equations Driven by Semimartingales," Stochastics, Vol. 4, pp. 223-245, 1981.

234. F. Konecny, "On Wong-Zakai Approximation of Stochastic Differential Equations," J. Multiv. Anal., Vol. 13, pp. 605-611, 1983.

235. R. R. Mohler and W. J. Kolodziej, "Optimal Control of a Class of Nonlinear Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 1048-1054, 1981.

236. R. M. Capocelli and L. M. Ricciardi, "On the Role of Noise in Systems Dynamics," Proc. Conf., pp. 1058-1064.

237. M. Athans, R. T. Ku and S. B. Gershwin, "The Uncertainty Threshold Principle: Some Fundamental Limitations of Optimal Decision Making Under Dynamic Uncertainty," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 491-495, 1977.

238. R. J. Ku and M. Athans, "Further Results on the Uncertainty Threshold Principle," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 866-868, 1977.

239. F. Kozin, "A Survey of Stability of Stochastic Systems," Automatica, Vol. 5, pp. 95-112, 1969.

240. D. L. Kleinman, "On the Stability of Linear Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 429-430, 1969.

241. J. C. Willems and G. L. Blankenship, "Frequency Domain Stability Criteria for Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 292-299, 1971.

242. J. L. Willems, "Mean Square Stability Criteria for Stochastic Feedback Systems," Int. J. Sys. Sci., Vol. 4, pp. 545-564, 1973.

243. J. L. Willems, "Mean Square Stability Criteria for Linear White Noise Stochastic Systems," Prob. Contr. Inf. Thy., Vol. 2, pp. 199-217, 1973.

244. U. G. Haussmann, "Stability of Linear Systems with Control Dependent Noise," SIAM J. Contr., Vol. 11, pp. 382-394, 1973.

245. U. Haussmann, "On the Existence of Moments of Stationary Linear Systems with Multiplicative Noise," SIAM J. Contr., Vol. 12, pp. 99-105, 1974.

246. A. S. Willsky, S. I. Marcus and D. N. Martin, "On the Stochastic Stability of Linear Systems Containing Colored Multiplicative Noise," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 711-713, 1975.

247. J. L. Willems and J. C. Willems, "Feedback Stabilizability for Stochastic Systems with State and Control Dependent Noise," Automatica, Vol. 12, pp. 277-283, 1976.

248. J. L. Williams and D. Aeyels, "An Equivalence Property for Moment Stability Criteria for Parametric Stochastic Systems and Ito Equations," Int. J. Sys. Sci., Vol. 7, pp. 577-590, 1976.

249. F. Kozin and S. Sugimoto, "Relations Between the Sample and Moment Stability for Linear Stochastic Differential Equations," in Proc. Conf. Stoch. Diff. Eqns., pp. 145-162, D. Mason, ed., Academic Press, New York, 1977.

250. J. L. Willems, "Moment Stability of Linear White Noise and Coloured Noise Systems," in Stochastic Problems in Dynamics, pp. 67-89, B. L. Clarkson, ed., Pitman, London, 1977.

251. A. Kistner, "On the Moments of Linear Systems Excited by a Coloured Noise Process," in Stochastic Problems in Dynamics, pp. 36-53, B. L. Clarkson, ed., Pitman, London, 1977.

252. T. Sasagawa, "On the Exponential Stability and Instability of Linear Stochastic Systems," Int. J. Contr., Vol. 33, pp. 363-370, 1981.

253. T. Sasagawa, "Sufficient Conditions for the Exponential p-Stability and p-Stabilizability of Linear Stochastic Systems," Int. J. Sys. Sci., Vol. 13, pp. 399-408, 1982.

254. Y. A. Phillis, "Entropy Stability of Continuous Dynamic Systems," Int. J. Contr., Vol. 35, pp. 323-340, 1982.

255. Y. A. Phillis, "Optimal Stabilization of Stochastic Systems," J. Math. Anal. Appl., Vol. 94, pp. 489-500, 1983.

256. J. L. Willems and J. C. Willems, "Robust Stabilization of Uncertain Systems," SIAM J. Contr. Optim., Vol. 21 pp. 352-374, 1983.

257. L. Arnold, H. Crauel and V. Wihstutz, "Stabilization of Linear Systems by Noise," SIAM J. Contr. Optim., Vol. 21, pp. 451-461, 1983.

258. R. F. Curtain and P. L. Falb, "Ito's Lemma in Infinite Dimensions," J. Math. Anal. Appl., Vol. 31, pp. 434-448, 1970.

259. R. F. Curtain, "Stochastic Differential Equations in Hilbert Space," J. Diff. Eqns., Vol. 10, pp. 412-430, 1971.

260. H. V. D. Water and J. C. Willems, "The Value of Information in Stochastic Control," RAIRO Autom. Sys. Anal. Contr., Vol. 17, pp. 113-129, 1983.

261. A. V. Balakrishnan, "Stochastic Bilinear Partial Differential Equations," Proc. 2nd USA-Italy Sem. Var. Str. Sys., May 1974, A. Ruberti, R. R. Mohler, eds., Springer-Verlag, 1975.

262. A. Germani and P. Sen, "White Noise Solutions for a Class of Distributed Feedback Systems with Multiplicative Noise," Ricerche di Autom., Vol. 10, pp. 38-65, 1979.

263. A. Gandolfi and A. Germani, "On the Definition of a Topology in Hilbert Space with Applications to the White Noise Theory," J. Franklin Inst., Vol. 316, pp. 435-444, 1983.

264. N. Berman and W. L. Root, "A Weak Stochastic Integral in Banach Space with Application to a Linear Stochastic Differential Equation," Appl. Math. Optim., Vol. 10, pp. 97-125, 1983.

265. K. Ito, Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces, Vol. 47, CBMS-NSF Req. Conf. Ser., SIAM, 1984.

266. C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Its Applications, John Wiley and Sons, New York, 1971.

267. S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

268. M. Athans, "The Matrix Minimum Principle," Inform. Contr., Vol. 11, pp. 592-606, 1968.

269. W. M. Wonham, Linear Multivariable Control: A Geometric Approach, Springer-Verlag, New York, 1974.

270. A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse," SIAM J. Appl. Math., Vol. 17, pp. 434-440, 1969.

271. E. Kreindler and A. Jameson, "Conditions for Nonnegativeness of Partitioned Matrices," IEEE Trans. Autom. Contr., Vol. AC-17, pp. 147-148, 1972.

272. S. Barnett and C. Storey, Matrix Methods in Stability Theory, Barnes and Noble, New York, 1976.

273. A. Graham, Kronecker Products and Matrix Calculus, Ellis Horwood, Chichester, 1981.

274. A. R. Tiedemann and W. L. DeKoning, "The Equivalent Discrete-Time Optimal Control Problem for Continuous-Time Systems with Stochastic Parameters," Int. J. Contr., Vol. 70, pp. 449-466, 1984.

275. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," IEEE Trans. Autom. Contr., Vol. AC-17, pp. 474-483, 1972.

276. A. Vinkler and L. J. Wood, "Multistep Guaranteed Cost Control of Linear Systems with Uncertain Parameters," J. Guid. Contr., Vol. 2, pp. 449-456, 1979.

277. A. K. Mahalanabis and R. Rana, "Guaranteed Cost Solution of Optimal Control and Game Problems for Uncertain Systems," Optim. Contr. Appl. Meth., Vol. 1, pp. 353-360, 1980.

278. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Wiley-Interscience, 1953.

279. H. J. Kushner, Approximation and Weak Convengence Methods for Random Processes with Applications to Stochastic Systems Theory, MIT Press, Cambridge, MA, 1984.

280. H. J. Kushner and H. Huang, "Limits for Parabolic Partial Differential Equations with Wide Band Stochastic Coefficients," Stochastics, to appear.

281. G. Blankenship and G. C. Papanicolaou, "Stability and Control of Stochastic Systems with Wide Band Noise Disturbances," SIAM J. Appl. Math., Vol. 34, pp. 437-476, 1978.

282. T. Hida, Brownian Motion, Springer-Verlag, 1980.

283. D. A. Dawson, "Stochastic Evolution Equations and Related Measure Processes," J. Multiv. Anal., Vol. 5, pp. 1-55.

284. A. Friedman, Stochastic Differential Equations and Applications, Vols. I, II, Academic Press, NY, 1975.

285. H. P. McKean, Stochastic Integrals, Academic Press, NY, 1969.

286. D. W. Stroock, "The Malliavin Calculus and its Application to Second Order Partial Differential Equations," Math. Sys. Thy., Vol. 14, pp. 141-171, 1981.

287. R. E. Skelton and P. C. Hughes, "Modal Cost Analysis for Linear Matrix Second-Order Systems," J. Dyn. Sys. Meas. Contr., Vol. 102, pp. 151-180, 1980.

288. I. M. Horowitz, Synthesis of Feedback Systems, Academic Press, New York, 1963.

289. D. G. Leunberger, Optimization by Vector Space Methods, John Wiley and Sons, New York, 1969.

290. R. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons, New York, 1970.

291. J. B. Cruz and W. R. Perkins, "A New Approach to the Sensitivity Problem in Multivariable Feedback System Design," IEEE Trans. Autom. Contr., Vol. AC-9, 1964.

292. M. G. Safonov and M. Athans, "Gain and Phase Margin for Multiloop LQG Regulators," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 173-179, 1977.

293. C. A. Harvey and G. Stein, "Quadratic Weights for Asymptotic Regulator Properties," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 378-387, 1978.

294. J. C. Doyle and G. Stein, "Robustness with Observers," IEEE Trans, Autom. Contr., Vol. AC-24, 1979.

295. N. K. Gupta, "Frequency-Shaped Cost Functionals: Extension of Linear-Quadratic-Gaussian Designs Methods," J. Guid. Contr., Vol. 3, 1980.

296. N. A. Lehtomaki, N. R. Sandell, Jr. and M. Athans, "Robustness Results in Linear-Quadratic Gaussian Based Multivariable Control Designs," IEEE Trans. Autom. Contr., Vol. AC-26, 1981.

297. J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 4-16, 1981.

298. G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 301-320, 1981.

299. G. Zames and B. A. Francis, "Feedback, Minimax Sensitivity, and Optimal Robustness," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 585-601, 1983.

300. B. A. Francis and G. Zames, "On H^{∞}-Optimal Sensitivity Theory for SISO Feedback Systems," IEEE Trans. Autom. Contr. Vol. AC-29, pp. 9-16, 1984.

301. B. A. Francis, J. W. Helton and G. Zames, "H^{∞}-Optimal Feedback Controllers for Linear Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 888-900, 1984.

302. I. Horowitz, "Quantitative Feedback Theory," Proc. IEEE, Vol. 129, pp. 215-226, 1982.

303. K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and their L^{∞}-Error Bounds," Int. J. Contr., Vol. 39, pp. 1115-1193, 1984.

304. D. G. Luenberger, "Observers for Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-11, 1966.

305. G. Menga and P. Dorato, "Observer-Feedback Design for Linear Systems with Large Parameter Uncertainties," Proc. IEEE Conf. Dec. Contr., 1974.

306. S. Gutman and G. Leitmann, "Stabilizing Feedback Control for Dynamical Systems with Bounded Uncertainty," Proc. IEEE Conf. Dec. Contr., 1976.

307. K. Furuta, S. Hara and S. Mori, "A Class of Systems with the Same Observer," IEEE Trans. Autom. Contr., Vol. AC-21, 1976.

308. R. V. Patel, M. Toda and B. Sridhar, "Robustness of Linear Quadratic State Feedback Designs in the Presence of System Uncertainty," IEEE Trans. Autom. Contr., Vol. AC-22, 1977.

309. J. B. Curz, Jr. and B. Krogh, "Design of Sensitivity-Reducing Compensators Using Observers," IEEE Trans. Autom. Contr., Vol. AC-23, 1978.

310. A. M. Meilakhs, "Design of Stable Control Systems Subject to Parametric Perturbation," Avtomatika i Telemekhanika, no. 10, 1978.

311. G. Leitmann, "Guaranteed Ultimate Boundedness for a Class of Uncertain Linear Dynamical Systems," IEEE Trans. Autom. Contr., Vol. AC-23, 1978.

312. G. Leitmann, "Guaranteed Asymptotic Stability for Some Linear Systems with Bounded Uncertainties," J. Dyn. Sys. Meas. Contr., Vol. 101, pp. 212-216, 1979.

313. S. Gutman, "Uncertain Dynamical Systems, A Lyapunov Min-Max Approach," IEEE Trans. Autom. Contr., Vol. AC-24, pp. 437-443, 1979.

314. M. Eslami and D. L. Russell, "On Stability with Large Parameter Variations: Stemming from the Direct Method of Lyapunov," IEEE Trans. Autom. Contr., Vol. AC-25, no. 6, 1980.

315. M. Corless and G. Leitmann, "Continuous State Feedback Guaranteeing Uniform Ultimate Boundedness for Uncertain Dynamical Systems," IEEE Trans. Autom. Contr., Vol. AC-26, 1139-1144, 1981.

316. J. S. Thorp and B. R. Barmish, "On Guaranteed Stability of Uncertain Linear Systems via Linear Control," J. Optim. Thy. Appl., Vol. 35, 1981.

317. B. R. Barmish and G. Leitmann, "On Ultimate Boundedness Control of Uncertain Systems in the Absence of Matching Conditions," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 153-157, 1982.

318. B. R. Barmish, M. Corless and G. Leitmann, "A New Class of Stabilizing Controllers for Uncertain Dynamical Systems," SIAM J. Contr. Optim., Vol. 21, pp. 246-255, 1983.

319. B. R. Barmish, "Stabilization of Uncertain Systems Via Linear Control," IEEE Trans. Autom. Contr., Vol. AC-28, no. 8, pp. 848-850, 1983.

320. R. T. Stefani, "Reducing the Sensitivity to Parameter Variations of a Minimum-Order Reduced-Order Observer," Int. J. Contr., Vol. 35, 1982.

321. G. Leitmann and W. Breinl, "State Feedback for Uncertain Dynamical Systems," Journ. Appl. Math. Comp.

322. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, Univ. of Illinois Press, Urbana, IL, 1949.

323. S. Guiash, Information Theory with Applications, McGraw-Hill, New York, 1974.

APPENDIX A

(REFERENCE [19])

COMPARISON OF VARIOUS CONTROLLER-REDUCTION METHODS: SUBOPTIMAL VERSUS OPTIMAL PROJECTION

84-1033

D. C. Hyland*

Harris Corporation, GASD, Melbourne, FL 32902

Abstract

Several suboptimal approaches to reduced-order controller design are compared with the new optimal projection formulation of the quadratically optimal fixed-form compensator problem. The substantial similarities and significant differences among the various design techniques are highlighted by placing the design equations of all methods within a common notation. Basically, all methods characterize the reduced-order controller by a projection on the full state space. The suboptimal methods construct this projection on the basis of balancing considerations while the optimal projection equations define it as a consequence of optimality conditions. Issues relating to relative computational simplicity and design reliability are explored by applying two of the methods to the same example problem.

1. Introduction

The design of reduced-order dynamic controllers for high-order systems is of considerable importance for applications involving large spacecraft and flexible flight systems and extensive research has recently been devoted to this area. This paper reviews and compares, both theoretically and computationally, several current approaches to reduced-order controller design.

One procedure for addressing this problem is to first apply some suitable model reduction algorithm to reduce the plant model to the dimension desired for the controller and then obtain an LQG controller which is optimal for the reduced model. A second, and perhaps more satisfactory approach is to predicate the control design upon a higher order model and then to reduce the dimension of the controller. Of course, this technique presupposes that some form of model reduction is still employed to reduce the originally very high order plant model to a "Riccati-solvable" dimension.

Because they all reflect the latter "controller reduction" philosophy and exhibit significant similarities, we confine attention here to the following methods:

1) Balanced Controller Reduction Algorithm (BCRA)

   --This is the internal balancing model reduction approach of Moore[1] applied to the controller reduction problem.

2) Balanced Controller Reduction Algorithm - Modified (BCRAM)

   --A modification of BCRA by Yousuff and Skelton.[2]

3) Component Cost Algorithm (CCA)

   --An application of Component Cost Analysis to this problem by Yousuff and Skelton.[3]

4) Optimal Projection Conditions (OPC)

   --These design equations are actually the necessary stationarity conditions for quadratically optimal, fixed-order compensator design in the form originally derived in Reference 4. Subsequently, the derivation was significantly improved and strengthened by Hyland and Bernstein.[5]

5) Approximate Optimal Projection Conditions -- nth iterate ($AOPC_n$)

   --This denotes the iterative algorithm, terminated at the nth iterate, for solution of the OPC as described in Reference 6. Under benign conditions, this algorithm is locally convergent so that $AOPC_n$ becomes equivalent to OPC as $n \to \infty$.

A basic distinction among the above methods should be noted: (1) - (3) are admittedly suboptimal approaches based, essentially, upon balancing considerations while formulation (4) and its computational implementation (5) arise from consideration of quadratically optimal, fixed-order compensator design.

Of course, the stationary conditions for fixed-form compensation have been written down (see [7-12], for example). However, full exploitation of the stationary conditions has no doubt been retarded by their extreme complexity. What is lacking, to quote the insightful remarks of [11], "is a deeper understanding of the structural coherence of these equations." The contribution of References [4] and [5] was to show how the orginally very complex stationary conditions can be reduced, without loss of generality, to much simpler and more tractable forms. The resulting equations preserve the simple form of LQG relations for the gains in terms of covariance and cost matrices, which, in turn, are determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations. This coupling, by means of a projection (idempotent matrix) whose rank is precisely equal to the order of the compensator, represents a graphic portrayal of the demise of the classical separation principle.

The compensator form which naturally emerges from this formulation is fully defined by the gains and by the projection matrix, whose row and column spaces are, respectively, the observation and control subspaces of the compensator. In fact, the stationary conditions are of such a form that they determine this "optimal projection" together with the gains.

It must be emphasized that the emergence of such a compensator projection does not represent an a priori assumption regarding the controller structure but rather is a consequence of the first-order necessary conditions.

---

* Leader, Control Systems Analysis and Synthesis Group, Member, AIAA.

The highly structured character of the optimal projection conditions not only gives rise to direct numerical solution procedures (as has been illustrated in [6]) but also sheds light on the various suboptimal techniques. One aim of this paper is to elucidate the fundamental connections existing between the fixed-order compensator optimality conditions and the balancing approaches of methods 1-3.

Note that since method 4 above constitutes the optimality conditions for fixed-order compensation, it theoretically represents the "best" controller-order reduction scheme--i.e. it gives the *minimum* (zero) degree of suboptimality and can thus serve as a standard of comparison. On the other hand, solution of OPC via the iterative algorithm of method 5 entails more computational effort than methods 1-3. Thus, the second major goal of this paper is to examine the tradeoffs between the greater computational simplicity of methods such as 1-3 versus the possibilities of improved performance and design reliability offered by method 5.

The plan of the paper is as follows. After presenting the general problem formulation, we establish a common notation and display the basic equations of all methods side by side (see Table 1 below). This permits the various design approaches to be compared quite directly and introduces considerable efficiency in the discussion. After addressing several important theoretical issues, we finally apply a selection of the methods to a single numerical example which involves a 20 state reduced-order version (see Reference [13]) of the CSDL ACOSS Model No. 2. The theoretical and numerical results allow several conslusions to be drawn regarding the comparative efficiency and suboptimality of the several design methods. In particular, it is shown that with modest increase in computational effort the optimal projection approach produces stable and optimal designs in cases wherein some suboptimal methods fail to yield stable designs.

2. Problem Statement - Setting Up a Common Notation

The problem addressed concerns the linear, finite-dimensional, time-invariant system:

$$\left.\begin{aligned} \dot{x} &= Ax + Bu + w_1; \quad X \in \mathbb{R}^N \\ Y &= Cx + w_2; \quad Y \in \mathbb{R}^P \end{aligned}\right\} \qquad (1)$$

where x is the plant state, A is the plant dynamics matrix and B and C are control input and sensor output maps, respectively. w_1 is a white disturbance noise with intensity matrix $V_1 \geq 0$ and w_2 is observation noise with nonsingular intensity $V_2 > 0$.

The problem addressed by all the methods listed above is to design a constant gain dynamic compensator of the form:

$$\left.\begin{aligned} u &= -Kq, \quad u \in \mathbb{R}^{\ell} \\ \dot{q} &= A_c q + FY, \quad q \in \mathbb{R}^{N_c} \\ N_c &< N \end{aligned}\right\} \qquad (2)$$

such that the quadratic, steady-state performance index:

$$\left.\begin{aligned} J_s &\triangleq \lim_{t_1 - t_0 \to \infty} J/|t_1 - t_0| \\ J &\triangleq \int_{t_0}^{t_1} dt \quad E[x^T R_1 x + u^T R_2 u] \\ R_1 &\geq 0, \quad R_2 > 0 \end{aligned}\right\} \qquad (3)$$

is either minimized (subject to the structural constraints implicit in (2)) or at least, rendered as small as is practicable. The challenging aspect of the problem is that in accordance with practical implementation constraints associated with the limitations of on-line software, N_c (the dimension of the compensator) is chosen in advance to be some number which is less than the plant dimension.

Within the above problem formulation, it is now possible to distill all methods considered here into a common notation. Although reasonably obvious, the assertions made below beginning with equation (4) and concluding with equation (16) and Table 1 are substantiated for methods 1-3 in the Appendix. No additional confirmation is required for the optimal projection equations since the following ideas were explicitly stated for OPC in [4] and [5].

First, it can be shown that all design methods considered establish a projection of rank N_c:

$$\tau \in \mathbb{R}^{N \times N}, \; \tau^2 = \tau, \; \text{rank}(\tau) = N_c \qquad (4)$$

which characterizes the observation and control subspaces encompassed by the compensator. Moreover, a factorization of τ is always employed (at least implicitly) which has the form:

$$\tau = g^T \Gamma \qquad (5.a)$$

where Γ and g are full-rank, N_c x N matrices satisfying:

$$\Gamma g^T = g \Gamma^T = I_{N_c} \qquad (5.b)$$

Note that if τ is a projection, then Γ and g satisfying (5.b) always exist such that (5.a) holds.

Secondly, it can be shown that methods 1-5 all produce matrices K, F, A_c of the form:

$$\left.\begin{aligned} K &= \hat{K} g^T \\ F &= \Gamma \hat{F} \\ A_c &= \Gamma (A - \hat{F}C - B\hat{K}) g^T \end{aligned}\right\} \qquad (6)$$

where

$$\left.\begin{aligned} \hat{K} &= R_2^{-1} B^T P \\ \hat{F} &= Q C^T V_2^{-1} \end{aligned}\right\} \qquad (7)$$

and where P and Q are both symetric and positive semi-definite, N X N matrices.

In summary, all methods yield the closed-loop system equations:

$$\left.\begin{aligned}\dot{x} &= Ax - B\hat{K}g^Tq + w_1\\ \dot{q} &= \Gamma(A - \hat{F}C - B\hat{K})g^Tq + \Gamma\hat{F}(CX = w_2)\end{aligned}\right\} \quad (8)$$

with $\hat{K}$ and $\hat{F}$ given by (7). In other words, the reduced-order compensator takes the form of a full-order compensator projected down to an N_C - dimensional subspace. The fact that (4) - (8) hold for the various suboptimal design methods was recently recognized in [14].

Thus, the principal distinctjon among the design methods rests in the manner in which P, Q and τ (or equivalently Γ and g) are constructed. To elucidate this matter, we first introduce the lemma (see [15]):

Lemma. *Suppose $M_1 \in \mathbb{R}^{NxN}$ and $M_2 \in \mathbb{R}^{NxN}$ are positive semi-definite. Then the product $M \triangleq M_1M_2$ is semi-simple (i.e. all Jordan blocks are of order unity) with real, non-negative eigenvalues.*

For convenience, let us also set up some additional notation relating to semi-simple matrices. If $M \in R^{NxN}$ is semi-simple, then, for some non-singular ϕ :

$$M = \phi\lambda\phi^{-1} \quad (9)$$

where λ is the diagonal matrix of eigenvalues of M.

Now, letting u_K denote the K^{th} column of ϕ and v_K^T the K^{th} row of ϕ^{-1}, (9) may be expressed as:

$$M = \sum_{k=1}^{N} \lambda_K u_K v_K^T \quad (10.a)$$

where the sets of vectors $\left|u_i\right|$, $\left|v_i\right|$ are mutually biorthonormal--i.e.:

$$v_K^T u_j = \begin{cases} 1 \; ; & K = j \\ 0 \; ; & K \neq j \end{cases} \quad (10.b)$$

and where we adopt the convention that the λ_K's are arranged in order of descreasing magnitude:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{N-1}| \geq |\lambda_N| \quad (10.c)$$

(10.a) is clearly analogous to the standard result for the spectral decomposition of a normal matrix. For this reason, we may term the quantify:

$$\pi_K[M] \triangleq u_K v_K^T \quad (11)$$

the eigen-projection of M associated with the K^{th} eigenvalue (under convention (10.c)). In view of (10.b), the $\pi_K[M]$ form a set of unit rank, mutually disjoint projections:

$$\left.\begin{aligned}(\pi_K[M])^2 &= \pi_K[M]\\ \pi_K[M]\,\pi_j\,[M] &= 0 \text{ if } K \neq j\end{aligned}\right\} \quad (12)$$

and M is written:

$$M = \sum_{K=1}^{N} \lambda_K \; \pi_K[M] \quad (13)$$

with convention (10.c).

This notation together with:

$$\Sigma \triangleq BR_2^{-1}B^T, \quad \bar{\Sigma} \triangleq C^TV_2^{-1}C \quad (14)$$

$$A_P \triangleq A - \Sigma P, \; A_Q \triangleq A - Q\bar{\Sigma}, \; \hat{A}_C \triangleq A - \Sigma P - Q\bar{\Sigma}$$

$$\hat{P} \triangleq P\Sigma P, \; \hat{Q} \triangleq Q\bar{\Sigma}Q \quad (15)$$

$$\tau_1 \triangleq I_N - \tau \quad (16)$$

allows us to state the basic design equations rather succinctly. Table 1 lists the equations determining P, Q and τ for BCRA, BCRAM, CCA, $AOPC_1$ and OPC, where P, Q, $\hat{P}$ and $\hat{Q}$ are required to be positive semi-definite. In view of the above Lemma, $\hat{Q}\hat{P}$ and $\hat{Q}\hat{P}$ are all semi-simple with real, non-negative eigenvalues. Thus all the methods displayed construct the projection τ as the sum of N_C (disjoint) eigenprojections associated with the N_C largest eigenvalues of a semi-simple matrix. Not shown in Table 1 is the computational algorithm, $AOPC_n$ for solution of the optimal projection equations. This algorithm proceeds as follows:

$AOPC_n$

1) To start, set $\tau_0 = I_N$
2) Using the previous iterate, τ_{K-1}, for τ, solve the OPC equations for P, Q, $\hat{P}$ and $\hat{Q}$.
3) Determine the eigenvalues and eigenvectors of $\hat{Q}\hat{P}$ and form the eigenprojections $\pi_K[\hat{Q}\hat{P}]$; K = 1,..., N.
(In general there will be $\tilde{N} > N_C$ non-zero eigenvalues. If there are exactly N_C non-zero eigenvalues at this point, then the OPC's are satisfied identically.)
4) Set τ_K equal to the sum of eigenprojections corresponding to the N_C largest eigenvalues of $\hat{Q}\hat{P}$.
5) Terminate if either (a) K = n or (b) ratio of the $(N_C + 1)^{th}$ to the $N_C{}^{th}$ eigenvalues of $\hat{Q}\hat{P}$ falls below some preassigned convergence tolerance, $\epsilon \ll 1$ (in which case the optimal projection conditions are satisfied to an acceptable approximation). Otherwise, increment K and return to Step 2.

In the following discussion, we shall constantly refer to Table 1 using the equation designations indicated--i.e. the first OPC equation will be referred to as Equation (OPC-a), etc.

Table 1

CCA	BCRA	BCRAM & $AOPC_1$	OPC	
$0 = PA + A^TP - \hat{p} + R_1$	$0 = PA + A^TP - \hat{p} + R_1$	$0 = PA + A^TP - \hat{p} + R_1$	$0 = PA + A^TP - \hat{p} + \tau_\perp^T \hat{p} \tau_\perp + R_1$	(a)
$0 = QA^T + AQ - \hat{q} + V_1$	$0 = QA^T + AQ - \hat{q} + V_1$	$0 = QA^T + AQ - \hat{q} + V_1$	$0 = QA^T + AQ - \hat{q} + \tau_\perp \hat{q} \tau_\perp^T + V_1$	(b)
$0 = \hat{Q}A_p^T + A_p \hat{Q} + \hat{q}$	$0 = \hat{Q}\hat{A}_c^T + \hat{A}_c\hat{Q} + \hat{q}$	$0 = \hat{Q}A_p^T + A_p\hat{Q} + \hat{q}$	$0 = \hat{Q}A_p^T + A_p\hat{Q} + \hat{q} - \tau_\perp \hat{q} \tau_\perp^T$	(c)
- - - - - -	$0 = \hat{P}\hat{A}_c + \hat{A}_c^T \hat{P} + \hat{p}$	$0 = \hat{P}A_Q + A_Q^T \hat{P} + \hat{p}$	$0 = \hat{P}A_Q + A_Q^T \hat{P} + \hat{p} - \tau_\perp^T \hat{p} \tau_\perp$	(d)
$\tau = \sum_{k=1}^{N_c} \Pi_k[\hat{Q}\hat{p}]$	$\tau = \sum_{k=1}^{N_c} \Pi_k[\hat{Q}\hat{P}]$	$\tau = \sum_{k=1}^{N_c} \Pi_k[\hat{Q}\hat{P}]$	$\tau = \sum_{k=1}^{N_c} \Pi_k[\hat{Q}\hat{P}]$	(e)

3. Theoretical Comparisons

The above description of $AOPC_n$ together with Table 1 summarizes all the design methods very succinctly. The similarities among the methods are evident. First, as already noted, all methods construct the compensator projection from the eigenprojections associated with the N_c largest eigenvalues of a product of two non-negative definite matrices. This provides additional motivation for the use of the term "optimal projection" in connection with the formulation of [4,5] since, there, the projection is determined via optimality not balancing considerations.

Secondly, all methods compute the cost matrices Q and P as solutions to Riccati equations or modified Riccati equations. Furthermore, BCRA, BCRAM and OPC construct τ from the product $\hat{Q}\hat{P}$ where $\hat{Q}$ and $\hat{P}$ are either controllability and observability grammians for the compensator (as in the case of BCRA) or are closely analogous quantities. The fact that CCA lacks a Lyapunov equation for $\hat{P}$ entails less of a distinction than might first be thought since the term $\hat{p}$ ($=P\Sigma P$) appearing in (CCA-e) is essentially the nonhomogeneous, driving term in the Lyapunov equations determining $\hat{P}$ in the other methods. Thus, the eigenvalues of $\hat{Q}\hat{p}$ (termed the "component costs" in [3]) assign a relative weighting to the eigenprojections in a manner analogous to the other methods.

In connection with equations (a) and (b) of the various methods, it was stated in [14] that the optimal projection design is simply the projection of an LQG controller. Equations (7) and (8) show that this would indeed be so if P and Q in (7) were determined as solutions to the LQG Riccati equations (as in CCA, BCRA and BCRAM). However, in contrast to the suboptimal approaches, the OPC equations for P and Q are <u>modified</u> Riccati equations containing <u>additional terms</u> involving the projection. Thus, except under very special circumstances, the optimal fixed-order compensator is generally <u>not</u> a projection of an LQG design.

Indeed, the one striking distinction is that in OPC, the equations determining P, Q, $\hat{P}$ and $\hat{Q}$ involve the compensator projeciton explicitly. In essence, the terms $\tau_\perp^T\hat{p}\tau_\perp$ and $\tau_\perp\hat{q}\tau_\perp^T$ (which are lacking in the suboptimal approaches) serve to couple the "model reduciton" portion of the problem (equations (OPC-C) and (OPC-d)) with the gain computation portion ((OPC-a) and (OPC-b)). An indirect result of this feature is that only OPC fully accounts for the fact that the loop is being closed by an N_c - order compensator. In fact, as mentioned previously, OPC constitutes the first-order necessary conditions for the optimization problem--the optimal fixed-order controller design <u>must</u> entail satisfaction of the optimal projection equations. Also, under mild geometric restrictions, solution of the optimal projection equations guarantees closed-loop stability. In view of these properties, OPC can serve as a theoretical standard of comparison for all the other (suboptimal) methods.

Judging from the appearance of the design equations, BCRAM is the one suboptimal method most similar to OPC. In fact, as indicated in Table 1, BCRAM and $AOPC_1$ are identical. Basically, the BCRAM equations are the optimal projection equations with the coupling terms in $\tau_\perp$ omitted, and numerical evidence presented in [2] suggests that BCRAM gives improved performance over BCRA (from which BCRAM was originally derived as a modification).

Note that the coupling terms $\tau_\perp^T\hat{p}\tau_\perp$ and $\tau_\perp\hat{q}\tau_\perp^T$ in OPC necessitate an iterative solution algorithm such as $AOPC_n$ while CCA, BCRA and BCRAM compute the compensator projection in only one step. Nevertheless, the suboptimal methods cannot subsequently improve τ if it happens to result in an unsuitable design (with poor performance or instability). $AOPC_n$, in contrast, offers the mechanism for progressive refinement of the design to achieve a controller which is as nearly optimal as desired.

A further issue of general importance is whether or not the various methods can produce a minimal order optimal compensator. In other words, there may exist a compensator of order M<N which yields the same performance as a full-order (N_c=N) compensator. It is highly desirable that the selected design method be capable of producing such a design when it exists. It turns out that all methods considered here meet this requirement. This is substantiated for BCRA, BCRAM in [1] and [2] respectively.

The capability of OPC to yield minimal order compensators follows generally from their status as optimality conditions. More specifically, however, if the LQG compensator has unobservable or uncontrollable poles, then the associated subspaces appear in the null space of $\hat{Q}\hat{P}$. This connection between unobservable/uncontrollable poles and the rank of $\hat{Q}\hat{P}$ was explored in [5]. Thus, if a minimal

realization of order M exists, then rank $[\hat{Q}\hat{P}] = M$ and setting $N_c = M$ in OPC gives the desired reduced order compensator.

To illustrate the point, consider the example given in [2]. In our notation, the defining matrices are:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = [1, -1] \tag{17}$$

$$V_1 = \begin{bmatrix} 1 & -1 \\ -1 & 16 \end{bmatrix}, \quad V_2 = 1 \tag{18}$$

$$R_1 = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}, \quad R_2 = 1 \tag{19}$$

The second-order, LQG compensator is:

$$\left.\begin{aligned} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} &= \begin{bmatrix} -3 & 0 \\ 5 & -4 \end{bmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} + \begin{pmatrix} 1 \\ -4 \end{pmatrix} Y \\ u &= -[2, 0]q \end{aligned}\right\} \tag{20}$$

It is clear from inspection of (20) that the minimal (unity order) compensator is obtained by deleting q_2 to get:

$$\left.\begin{aligned} \dot{q} &= -3q + y \\ u &= -2q \end{aligned}\right\} \tag{21}$$

At the same time, it is easily verified that the optimal projection equations yield the solution:

$$\left.\begin{aligned} P &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix} \\ \beta &= 17/4 + 16/17 \\ \hat{P} &= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{Q} = \frac{17}{4} \begin{bmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{bmatrix} \\ \alpha &= -32/17 \end{aligned}\right\} \tag{22}$$

and

$$\tau = \begin{bmatrix} 1 & 0 \\ \alpha & 0 \end{bmatrix} \tag{23}$$

or

$$\Gamma = [1, 0], \quad g = [1, \alpha] \tag{24}$$

Using these results with (6) and (7) shows that $K = 2$, $F = 1$ and $A_c = -3$. Thus, OPC yields the minimal order compensator, (21). When one applies the iterative solution scheme, AOPC$_n$, it is found that the correct projection and the desired values of K, F and A_c are produced on the first iteration. Further iterations beyond AOPC$_1$ yield no change in τ, K, F and A_c. Incidentally this also illustrates how both AOPC$_1$ <u>and</u> BCRAM yield the minimal order compensator.

At this point, it should be mentioned that the new "LQG - balancing" method of Verriest [16, 17] and Jonckheere and Silverman [18] will <u>not</u> always yield a minimal order compensator when <u>it</u> exists (and the property was, in fact, never claimed by the authors). In view of this distinction, and because the LQG - balancing method has been evaluated and compared extensively in [2] and [3], it is not given detailed consideration in the present paper.

4. Numerical Comparison Using A Common Example Problem

Finally, we explore issues relating to practical design efficiency by applying two of the methods considered here to the same example problem. We considered pointing and shape control of the "Solar Optical Telescope" spacecraft example discussed in [3]. The original 44 mode model was reduced to 10 modes (8 elastic and 2 rigid-body modes) by a Modal Cost Analysis in [13]. In the notation used here, the matrices defining this 20-state problem are:

$$\left.\begin{aligned} A &= \begin{bmatrix} 0 & I_{10} \\ -\omega^2 & -2\xi\omega \end{bmatrix} \\ \xi &= 0.001 \\ \omega &= \operatorname{diag}_{k=1,\dots,20} [\omega_k] \end{aligned}\right\} \tag{25}$$

$$B = \begin{bmatrix} 0 \\ \beta \end{bmatrix}, \quad C = [\,P, 0] \tag{26}$$

$$\left.\begin{aligned} V_1 &= 10^{-4} \begin{bmatrix} 0 \\ \beta \end{bmatrix} [0\ \beta^T] \\ V_2 &= 10^{-15}\ I_3 \end{aligned}\right\} \tag{27}$$

$$\left.\begin{aligned} R_1 &= \begin{bmatrix} P^T \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10^{-3} \end{bmatrix} [\,P, 0] && \text{a.} \\ R_2 &= \rho I_8 && \text{b.} \end{aligned}\right\} \tag{28}$$

where the modal frequencies, (ω_k, k = 1,..., 10) and matrices β and P are given in Table 1 and 2, respectively, of [3]. In (28.b), ρ is a positive scalar used to adjust the relative weighting of the state and control input penalties of the performance index. Clearly, overall controller authority, actuator mean-square force levels and compensator bandwidth are all inversely proportional to ρ.

Here, we discuss numerical results for $\rho \in [0.01, 100.0]$ and for N_c in the range from 20 to 4 for design methods CCA and AOPC$_n$. The approach adopted for design comparison is to plot "regulation cost" ($E[x^T R_1 x]$) as a function of "control cost" ($E[u^T u]$) (obtained by varying ρ) for each value of N_c and for each of the design methods.

Results for these tradeoff curves are shown in Fig. 1. The very bottom-most curve represents the full-order (20 states) LQG design. Since this is the best obtainable when there is no restriction on compensator order, the problem is to obtain a lower order design whose tradeoff curve is as close to the LQG results as possible.

The thin black lines in Figure 1 show the N_c=10, 6, and 4 designs obtained via Component Cost Analysis, where N_c denotes the compensator dimension. These results were obtained in [3] using the

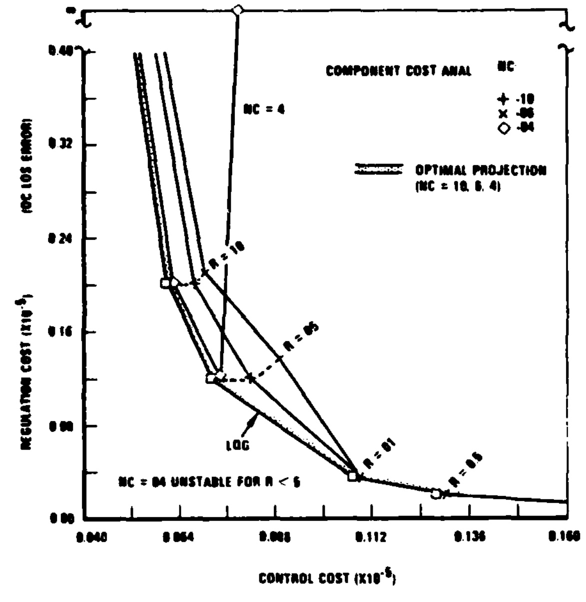


Fig. 1 Performance Tradeoff Curves For Component Cost Analysis and Optimal Projection

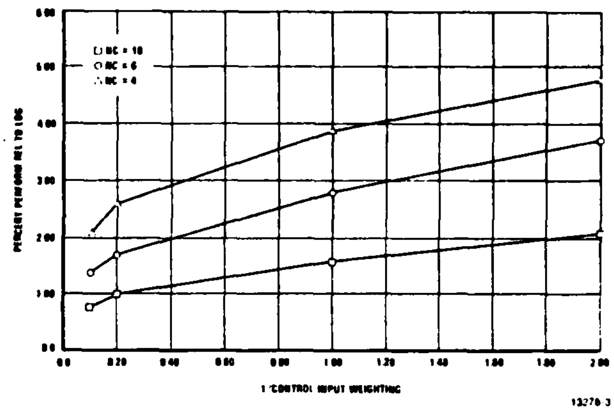


Fig. 2 Optimal Projection Results for S.O.T. Example - Percent Performance Increase Over LQG Design

full design algorithm described in Appendix A, [3] - including the refinements of steps III.a through III.e. Note that the 10th and 6th order compensator designs are quite good, but when compensator order is sufficiently low (N_c=4) and controller bandwidth sufficiently large (ρ<5.0), the method fails to yield stable designs. This difficulty is characteristic of suboptimal techniques, and in fairness, it should be noted that other suboptimal design methods (such as the LQG balanced design method proposed by Verriest [16, 17]) fail to give stable designs for compensator orders below 10.

In contrast, the width of the grey line in Figure 1 encompasses all the optimal projection results for compensators of order 10, 6, and 4. To provide a more detailed picture of the optimal projection results, Figure 2 shows the percent of total performance increase relative to the full-order, LQG designs (the quantity (100 x ($J_s(N_c)$-$J_s(20)$)/ $J_s(20)$) as a function of 1/ρ (proportional to controller bandwidth and to actuator force levels) for the various compensator orders considered.

Even for the 4th order design, the optimal projection performance is only ~5 percent higher than the optimal full-order design. Furthermore, the performance index for the optimal projection designs increases montonically with decreasing controller order - as it should. Such is not necessarily the case for suboptimal design methods.

The OPC results were actually obtained via the iterative algorithm $AOPC_n$ outlined in Section 2 (and described in more detail in [6]). In all cases except NC = 4, ρ = 0.5, adequate convergence in performance was obtained in four iterations. For the case NC = 4, ρ = 0.5, four iterations produced an unstable closed-loop system. This is apparently due to inadequate convergence in this case, since four additional iterations sufficed to give the results indicated in Fig's 1-2.

Note from Table 1 that one iteration of $AOPC_n$ requires only slightly more computation than CCA since CCA involves solution of only one Lyapunov equation and lacks an analogous equation for $\hat{P}$. Thus, it may be estimated that the OPC results given here required rougly 4 to 8 times the computational effort of CCA. This increased computational burden for OPC is offset, in the higher bandwidth cases, by the reliable production of very low-order but excellent performance designs.

5. Conclusions

In the preceeding, we have established a common theoretical framework within which various "balancing" approaches (BCRA, BCRAM and CCA) to controller order reduction may be compared directly with a new formulation (OPC) of the quadratically optimal fixed-order compensation problem. The optimal and suboptimal approaches were found to be highly analogous, each technique characterizing the reduced-order compensator by a projection of rank equal to the desired compensator dimension. The major distinction among the methods is that in OPC, the projection arises as a consequence of optimality conditions, while in the suboptimal methods it is constructed on the basis of other considerations. The ramifications of this distinction were explored by comparing performance results obtained via CCA and OPC for the same example problem.

The above comparisons lead to the following conclusions:

1) In all cases considered, the total performance index for OPC is less than for CCA. Even in the case N_c = 4, the OPC regulation vs. control cost plot is very close to that of the full-order (N_c = 20) design.

2) The performance index for OPC increases monotonically with decreasing controller order. Such is not generally the case for

suboptimal design methods.

3) Various suboptimal methods (including CCA) fail to produce stable designs when the compensator order and control input weighting parameter, ρ , are both sufficiently small. Contrasting results with OPC solutions, we see that this effect is often an artifact of the suboptimal character of the design methods and does not necessarily imply non-existance of stabilizing compensator designs for small N_C.

4) The computational effort associated with the use of CCA, BCRA or BCRAM is roughly equal to the effort of one iteration of the $AOPC_n$ algorithm used for solution of the optimal projection conditions. On the other hand, excellent results for OPC were obtained in 4 - 8 iterations. Thus, while it is significantly more reliable in producing satisfactory designs, OPC does necessitate an increase in computational burden.

The above remarks would seem to motivate continued exploration of the optimal projection approach and additional efforts to increase its computational efficiency. It is also possible that the optimal projection equations may provide the insight needed to devise improved suboptimal (and non-iterative) design techniques.

APPENDIX

Here we verify the statements made in Section 2. from equation (4) through (16) and Table 1 for the design methods *BCRA, BCRAM and CCA.*

BCRA

Here, consider the internal balancing approach to model reduction as applied to a full-order LQG compensator:

$$\left.\begin{aligned} \dot{q} &= A_C q + FY \\ u &= -Kq \end{aligned}\right\} \quad (A.1)$$

where A_C, K and F are given by (6) and (7) with $g = \Gamma = I_N$ and P and Q satisfying the LQG Riccati equations. Considering $V_2^{-\frac{1}{2}} y$ as the system input and $R_2^{\frac{1}{2}} u$ as the output, and assuming, in this discussion that A_C is asymptotically stable, the controllability and observability Grammians are determined as unique positive semidefinite solutions to (see Ref. [1], p. 21 and make allowance for notation and input, output definitions).

$$\left.\begin{aligned} 0 &= A_C W_C^2 + W_C^2 A_C^T + FV_2F^T \\ 0 &= A_C^T W_O^2 + W_O^2 A_C + K^T R_2 K \end{aligned}\right\} \quad (A.2)$$

Having obtained these quantities, it is clear from the discussion of [1], p. 24, or from the Lemma in the main text of this paper that there exists a transformation with matrix P such that W_C and W_O can both be reduced to the diagonal matrix of 2nd order modes, (Σ^2 in Moore's notation). Referring to the expressions given for W_C^2 and W_O^2 under coordinate transformation given in [1], p. 23, we have

$$\left.\begin{aligned} W_C^2(P) &= P^{-1} \Sigma^2 P^{-1T}, \\ W_O^2(P) &= P^T \Sigma^2 P \end{aligned}\right\} \quad (A.3)$$

Obviously, (A.3) indicates that P^{-1} is the right eigenvector matrix and Σ^4, the eigenvalues of $W_C^2W_O^2$. Thus, employing the notation introduced in section 2:

$$W_C^2 W_O^2 = \sum_{k=1}^{N} \Sigma_K^4 \, \pi_K[W_C^2 W_O^2] \quad (A.4)$$

where the π_k are formed directly from the columns and rows of P^{-1} and P, respectively. As recommended in the Internal Dominance section of [1], one forms a reduced-order model of (A.1) by deleting, in the internally balanced basis, all states associated with the $N - N_C$ smallest second order modes. Setting:

$$\left.\begin{aligned} \Gamma &= 1^{st}\ N_C \text{ rows of } P \\ g^T &= 1^{st}\ N_C \text{ columns of } P^{-1} \end{aligned}\right\} \quad (A.5)$$

this is tantamount to defining a reduced order model of (A.1) via (6) and (7). Thus, (6) and (7) are verified for BCRA.

Now Γ and g as obtained above are obviously one factorization of the projection:

$$\begin{aligned} \tau &= g^T\Gamma \\ &= P^{-1} \begin{bmatrix} I_{N_C} & 0 \\ 0 & 0 \end{bmatrix} P \\ &= \sum_{k=1}^{N_C} \pi_k[W_C^2 W_O^2] \end{aligned} \quad (A.6)$$

However, any factorization of τ can be related to a given one by a similarity transformation of the reduced order system. Since the second-order modes and the compensator performance are invariant under such transformations, the particular factorization is immaterial. Thus the internal balancing approach to controller reduction is mathematically equivalent to defining a controller of the form (6) and (7) by (1) computing P and Q via LQG Riccati equations (BCRA-a and BCRA-b of Table 1) (2) determining $\hat{Q} = W_C^2$ and $\hat{P} = W_O^2$ from equations (A.2) (setting $\hat{Q} = W_C^2$, $\hat{P} = W_O^2$, $F = QC^TV_2^{-1}$, $K = R_2^{-1}B^TP$ and and $\hat{A}_C = A - \Sigma P - Q\bar{\Sigma}$, equations (A.2) become BCRA-C and BCRA-d), (3) forming the compensator projection via (A.6) or, equivalently, (BCRA-e), and then effecting any suitable factorization in accordance with (5.a, b). This verifies (4) through (16) and Table 1 for BCRA.

BCRAM

This design method is a modification of BCRA proposed in [2] to circumvent difficulties in the application of BCRA when A_C is not asymptotically stable. Referring to equations (2.7.a) and (2.7.b) of [2], the only change from BCRA is that in (BCRA-c), $\hat{A}_C$ is replaced by $A - BK = A - \Sigma P$ and in (BCRA-d), $\hat{A}_C$ is replaced by $A - FC = A - Q\bar{\Sigma}$. Note that since P and Q are LQG Riccati equation solutions, $A - \Sigma P$ and $A - Q\bar{\Sigma}$ are asymptotically stable so that the method has a wider field of application.

In view of the preceding discussion, of BCRA, the simple changes noted suffice to substantiate (4) through (16) and Table 1 for BCRAM.

CCA

For simplicity, we consider only the basic Component Cost Analysis approach (termed the Cost-Decoupled Controller Design Algorithm) given in [3], omitting the special procedures described after Theorem 1 of Section III, Ref. [3] for identifying "nearly" unobservable states by use of a set of cost decoupled coordinates closely associated with a generalized Hessenburg representation. This means, that we confine attention to the algorithm in Appendix A of [3] without the computational refinements of steps III and IV.

Referring now to Appendix A, Ref. [3], the Cost-Decoupled design algorithm starts in steps Ia-Ib by computing the full-order, LQG compensator design. Allowing for differences in notation, the expressions (A.3b) - (A.3d) of Ref. [3] are identical to (6) and (7) with $\Gamma = g = I_N$. Similarly (A.3e) and (A.3f) of Ref. [3] are our equations (CCA-a) and (CCA-b) in Table 1.

Next consider step I.C. The quantities $\hat{X}$, BG and FVF^T in [3] correspond to $\hat{Q}$, $-\Sigma P$ and $Q\bar{\Sigma}Q$ in our notation. Thus, equaiton (A.4) of Ref. [3] is equivalent to our (CCA-c) for determination of $\hat{Q}$. Noting that G^TRG of [3] corresponds to the quantity $P\Sigma P$, steps I.d through II.a define a transformation matrix T_1:

$$T_1 = \Theta_x \Theta_u \tag{A.7}$$

where Θ_x *is the square root of* $\hat{Q}$:

$$\hat{Q} = \Theta_x \Theta_x^T \tag{A.8}$$

and Θ_u is the orthonormal modal matrix of $\Theta_x^T P\Sigma P\Theta_x$:

$$\left.\begin{aligned} \Theta_u^T \Theta_x^T P\Sigma P\Theta_x\Theta_u &= \operatorname{diag}_k \left|\nu_k\right| \\ \nu_1 \geq \nu_2 \geq \cdots &\geq \nu_N \geq 0 \end{aligned}\right\} \tag{A.9}$$

where the ν_k's are the nonnegative "component costs".

Some simple manipulation of (A.7) - (A.9) shows that T_1 is defined such that:

$$\left.\begin{aligned} (\hat{Q}P\Sigma P)T_1 &= T_1 \operatorname{diag}_k \left|\nu_k\right| \\ T_1^{-1}\hat{Q}T_1^{-1T} &= I_N \\ T_1^T P\Sigma PT_1 &= \operatorname{diag}_k \left|\nu_k\right| \end{aligned}\right\} \tag{A.10}$$

-i.e. T_1 is a right eigenvector matrix of the product $\hat{Q}(P\Sigma P)$ yielding a state transformation which simultaneously diagonalizes the factors $\hat{Q}$ and $P\Sigma P$.

Step III basically effects a refinement of T_1 to isolate weakly observable components of the state. For convenience in the exposition we omit this here to obtain a statement of the "basic" CCA algorithm. Thus, set $T = T_1$ in Step IV and proceed to Step V. Equations (A.13e) and (A.13f) of [3] *read:*

$$\left.\begin{aligned} [T_R,\ T_T] &\triangleq T_1 \quad ; \quad T_R \in \mathbb{R}^{N\times N_c} \\ \begin{bmatrix} L_R \\ L_T \end{bmatrix} &\triangleq T_1^{-1} \quad ; \quad L_R \in \mathbb{R}^{N_c\times N} \end{aligned}\right\} \tag{A.11}$$

Clearly $T_R L_R$ defines a rank N_c projection on $\mathbb{R}^N$ which, in view of (A.10), is the sum of the eigenprojections of $\hat{Q}P\Sigma P$ associated with the N_c largest eigenvectors. This verifies (CCA-e) of Table 1. Also, comparing $\tau = T_R L_R$ with (5.a), it is seen that T_R and L_R correspond to g^T and Γ, respectively. Thus, allowing for notational changes, equations (A.13b - A.13d) of [3] read:

$$\left.\begin{aligned} A_c &= \Gamma(A - Q\bar{\Sigma} - \Sigma P)g^T \\ F &= \Gamma QC^T V_2^{-1} \\ K &= R_2^{-1} B^T P \end{aligned}\right\} \tag{A.12}$$

where P and Q are solutions to (CCA-a) and (CCA-b). (A-12) are obviously equivalent to (6) and (7).

REFERENCES

1. B.D. Moore; "Principal Component Analysis in Linear Systems; Controllability, Observability and Model Reduction", IEEE Trans. Automat. Contr., Vol. AC-26, pp. 17-32, 1981.

2. A. Yousuff and R. E. Skelton, "A Note on Balanced Controller Reduction", IEEE Trans. Automat. Contr., 1983 - to appear.

3. A. Yousuff and R. E. Skelton, "Controller Reduction by Component Cost Analysis", IEEE Trans. Automat. Contr., 1984 - to appear.

4. D. C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters", AIAA 20th Aerospace Sciences Mtg., Orlando, Florida, January, 1982.

5. D. C. Hyland and D. S. Bernstein, "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation," IEEE 22nd CDC, San Antonio, Texas, Dec. 14-16, 1983.

6. D. C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results", AIAA 21st Aerospace Sciences Mtg., Reno, Nevada, Jan. 9183.

7. C. J. Wenk and C. H. Knapp, "Parameter Optimization in Linear Systems with Arbitrarily Constrained Controller Structure", IEEE Transactions on Automatic Control, Vol. AC-25, pp. 496-500, 1980.

8. W. S. Levine, T. L. Johnson and M. Athans, "Optimal Limited State Variable Feedback Controllers for Linear Systems", IEEE Transactions on Automatic control, Vol. AC-16, pp. 785-793, 1971.

9. T. L. Johnson and M. Athans, "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems", IEEE Transactions on Automatic Control, Vol. 15, pp. 658-660, 1970.

10. R. B. Asher and J. C. Durrett, "Linear Discrete Stochastic Control with a Reduced-Order Dynamic Compensator" IEEE Transactions on Automatic Control, Vol. AC-21, pp. 626-627, 1976.

11. W. J. Naeije and O. H. Bosgra, "The Design of Dynamic Compensators for Linear Multivariable Systems", 1977 IFAC, Fredricksburg, Canada, pp. 205-212.

12. S. Basuthakur and C. H. Knapp, "Optimal Constant Controllers for Stochastic Linear Systems", IEEE Transactions on Automatic Control. Vol. AC-20, pp. 664-666, 1975.

13. R. E. Skelton, P. C. Hughes, "Modal Cost Analysis for Linear Matrix Second-Order Systems, J. Dyn. Syst. Meas. and Control, Vol. 102, Sept. 1980, pp. 151-180.

14. A. Yousuff, R. E. Skelton, "A Projection Approach to Controller Reduction", Proc. IEEE 22nd CDC, San Antonio, Texas, Dec., 1983.

15. C. R. Rao and S. H. Mitra, Generalized Inverse of Matrices and Its Application, John Wiley & Sons, Inc., New York, 1971, p. 123, Theorem 6.2.5.

16. E. I. Verriest, "Low Sensitivity Design and Optimal Order Reduction for the LQG Problem", Proc. 24th Symp. Circuits and Systems, pp. 365-369, June 1981.

17. E. I. Verriest, "Suboptimal LQG-design and Balanced Realizations", Proc. 20th IEEE CDC, pp. 686-687, Dec. 1981.

18. E. A. Jonckheere, L. M. Silverman, "A New Set of Invariants for Linear Systems - Application to Approximation", International Symposium Math. Th. Networks and Syst., Santa Monica, CA, 1981.

APPENDIX B

(REFERENCE [23])

The Optimal Projection Equations for Fixed-Order Dynamic Compensation

DAVID C. HYLAND AND DENNIS S. BERNSTEIN

Abstract—**First-order necessary conditions for quadratically optimal, steady-state, fixed-order dynamic compensation of a linear, time-invariant plant in the presence of disturbance and observation noise are derived in a new and highly simplified form. In contrast to the pair of matrix Riccati equations for the full-order LQG case, the optimal steady-state fixed-order dynamic compensator is characterized by *four* matrix equations (two modified Riccati equations and two modified Lyapunov equations) coupled by a projection whose rank is precisely equal to the order of the compensator and which determines the optimal compensator gains. The coupling represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case.**

I. INTRODUCTION

Because of constraints imposed by on-line computations, dynamic controllers for high-order systems such as flexible spacecraft must be of relatively modest order. Hence, this paper is concerned with the design of quadratically optimal, fixed-order (i.e., reduced-order) dynamic compensation for a plant subject to stochastic disturbances and nonsingular measurement noise. Since white noise in all measurement channels precludes direct output feedback (see Section II), only purely dynamic controllers are considered. The requirements for resolution of this optimization problem include the following.

1) Conditions for the existence of an optimal, stabilizing compensator of the prescribed order. (In the full-order case these are the usual stabilizability and detectability conditions of LQG theory.)

2) Stationary conditions, i.e., first-order necessary conditions, rendered in a tractable form to facilitate developments in items 3) and 4) below. (In the full-order case these conditions are precisely the LQG gain relations together with the regulator and observer Riccati equations.)

3) Sufficiency conditions, i.e., additional restrictions on solutions of the first-order necessary conditions which characterize local minima and single out the *global* minimum. (In the full-order case the global minimum is distinguished by the unique nonnegative-definite solutions to the LQG Riccati equations.)

4) Convergent numerical algorithms for simultaneous satisfaction of the necessary and sufficient conditions. (In the full-order case numerical algorithms have been devised which take full advantage of the highly structured form of the Riccati equations.)

Manuscript received August 15, 1983; revised February 14, 1983. This paper is based on a prior submission of March 10, 1983. Paper recommended by P. R. Kumar, Past Chairman of the Stochastic Control Committee. This work was supported by the Department of the Air Force and was performed at Lincoln Laboratory, M.I.T.

The authors are with the Controls Analysis and Synthesis Group, Harris Corp., GASD, Melbourne, FL 32902.

0018-9286/84/1100-1034$01.00

The present paper deals exclusively with item 2). Although the stationary conditions for the fixed-order compensation problem *have* been written down (see [1]–[12], for example), full exploitation has undoubtedly been impeded by their extreme complexity [see (3.3)–(3.11)]. What has been lacking, to quote the insightful remarks of [9], "is a deeper understanding of the structural coherence of these equations." The contribution of the present paper is to show how the originally *very* complex stationary conditions can be transformed without loss of generality to much simpler and more tractable forms. The resulting equations (2.10)–(2.17) preserve the simple form of LQG relations for the gains in terms of covariance and cost matrices which, in turn, are determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations. This coupling, by means of a projection (idempotent matrix) whose rank is pecisely equal to the order of the compensator, represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case. When, as a special case, the order of the compensator is required to be equal to the order of the plant, the modified Riccati equations reduce to the standard LQG Riccati equations and the modified Lyapunov equations express the proviso that the compensator be minimal, i.e., controllable and observable. Since the LQG Riccati equations as such are nothing more than the necessary conditions for full-order compensation, we believe that the "optimal projection equations" provide a clear and simple generalization of standard LQG theory.

Since we are concerned with *optimal* fixed-order compensator design, our approach does not represent yet another model- or controller-reduction scheme along the lines of [13]–[17]. Indeed, the optimal projection equations, by virtue of their relatively transparent structure, can reveal the extent to which the design equations of a given ad hoc reduction scheme conform to the necessary conditions for optimality. For example, the oblique projection which arises in the present formulation may not be of the form $\left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right]$ even in the basis corresponding to the "balanced" realization [13]–[16]. These issues are discussed in [18] where the results of [19] are simplified by means of the approach of the present paper and where the balancing method of [13] is reinterpreted in the context of optimality theory.

The fact that the optimal projection equations consist of four coupled matrix equations, i.e., two modified Riccati equations and two modified Lyapunov equations, should not be at all surprising for the following simple reason. Reduced-order control-design methods often involve either LQG applied to a reduced-order model or model reduction applied to a full-order LQG design. Both approaches, then, involve the solution of precisely four equations: two Riccati equations (for LQG) plus two Lyapunov equations (for model reduction via balancing, as in [13]). The *coupled* form of the optimal projection equations is thus a strong reminder that the LQG and order-reduction operations *cannot* be iterated but must, in a certain sense, be performed simultaneously.

II. Problem Statement and the Main Theorem

Given the control system

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t), \tag{2.1}$$

$$y(t) = Cx(t) + w_2(t) \tag{2.2}$$

design a fixed-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \tag{2.3}$$

$$u(t) = C_c x_c(t) \tag{2.4}$$

which minimizes the steady-state performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t\to\infty} \mathbb{E}[x(t)^T R_1 x(t) + u(t)^T R_2 u(t)] \tag{2.5}$$

where: $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^l$, $x_c \in \mathbb{R}^{n_c}$, $n_c \leq n$, $A, B, C, A_c, B_c, C_c, R_1$, and R_2 are matrices of appropriate dimension with R_1 (symmetric) nonnegative definite and R_2 (symmetric) positive definite; w_1 is white disturbance noise with $n \times n$ nonnegative-definite intensity V_1 and w_2 is white observation noise with $l \times l$ positive-definite intensity V_2; w_1 and w_2 are uncorrelated and have zero mean. We note that the assumptions of nonsingular control weighting and nonsingular observation noise preclude the use of direct output feedback as in

$$u(t) = C_c x_c(t) + D_c y(t) \tag{2.6}$$

since J is undefined unless (see [7])

$$\mathrm{tr}\,[D_c^T R_2 D_c V_2] = 0 \quad (\Leftrightarrow R_2 D_c V_2 = 0). \tag{2.7}$$

To guarantee that J is finite and independent of initial conditions we restrict our attention to the set of admissible stabilizing compensators

$$\mathfrak{A} \triangleq \left\{ (A_c, B_c, C_c)\colon \tilde{A} \triangleq \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix} \text{ is asymptotically stable} \right\}$$

where $\tilde{A}$ is the closed-loop dynamics matrix. Since the value of J is independent of the internal realization of the compensator, we can further restrict our attention to

$$\mathfrak{A}_+ \triangleq \{(A_c, B_c, C_c) \in \mathfrak{A}\colon$$

$$(A_c, B_c) \text{ is controllable and } (C_c, A_c) \text{ is observable}\}.$$

For the following lemma call a square matrix nonnegative (respectively, positive) semisimple if it has a diagonal Jordan form and nonnegative (respectively, positive) eigenvalues. Let I_r denote the $r \times r$ identity matrix.

Lemma 2.1: Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n\times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if rank $\hat{Q}\hat{P} = n_c$ then there exist $G, \Gamma \in \mathbb{R}^{n_c \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \tag{2.8}$$

$$\Gamma G^T = I_{n_c}. \tag{2.9}$$

Proof. The result is an immediate consequence of [20, Theorem 6.2.5, p. 123]. □

For convenience in stating the Main Theorem, define

$$\Sigma \triangleq BR_2^{-1}B^T,\ \bar{\Sigma} \triangleq C^T V_2^{-1} C$$

and call G, M, and Γ satisfying (2.8) and (2.9) a (G, M, Γ)-factorization of $\hat{Q}\hat{P}$.

Main Theorem: Suppose $(A_c, B_c, C_c) \in \mathfrak{A}_+$ solves the steady-state fixed-order dynamic-compensation problem. Then there exist $n \times n$ nonnegative-definite matrices Q, P, $\hat{Q}$, and $\hat{P}$ such that A_c, B_c, and C_c are given by

$$A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P)G^T, \tag{2.10}$$

$$B_c = \Gamma Q C^T V_2^{-1}, \tag{2.11}$$

$$C_c = -R_2^{-1} B^T P G^T \tag{2.12}$$

for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$, and such that with $\tau \triangleq G^T\Gamma$ the following conditions are satisfied:

$$0 = (A - \tau Q\bar{\Sigma})Q + Q(A - \tau Q\bar{\Sigma})^T + V_1 + \tau Q\bar{\Sigma} Q \tau^T, \tag{2.13}$$

$$0 = (A - \Sigma P\tau)^T P + P(A - \Sigma P\tau) + R_1 + \tau^T P\Sigma P\tau, \tag{2.14}$$

$$0 = \tau[(A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^T + Q\bar{\Sigma} Q], \tag{2.15}$$

$$0 = [(A - Q\bar{\Sigma})^T\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P]\tau, \tag{2.16}$$

$$\mathrm{rank}\ \hat{Q} = \mathrm{rank}\ \hat{P} = \mathrm{rank}\ \hat{Q}\hat{P} = n_c. \tag{2.17}$$

Remark 2.1: Because of (2.9) the $n \times n$ matrix τ which couples the four equations (2.13)–(2.16) is idempotent, i.e., $\tau^2 = \tau$. In general this "optimal projection" is an *oblique* projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Note that Sylvester's inequality and (2.9) imply that rank $\tau = n_c$.

Remark 2.2: Using the relations $\hat{Q} = \tau\hat{Q}$ and $\hat{P} = \hat{P}\tau$ [see (3.12)],

the optimal projection equations (2.13)–(2.16) can be written in the equivalent form

$$0=AQ+QA^T+V_1-Q\Sigma Q+\tau_\perp Q\Sigma Q\tau_\perp^T, \tag{2.18}$$

$$0=A^TP+PA+R_1-P\Sigma P+\tau_\perp^T P\Sigma P\tau_\perp, \tag{2.19}$$

$$0=(A-\Sigma P)\hat{Q}+\hat{Q}(A-\Sigma P)^T+Q\Sigma Q-\tau_\perp Q\Sigma Q\tau_\perp^T, \tag{2.20}$$

$$0=(A-Q\Sigma)^T\hat{P}+\hat{P}(A-Q\Sigma)+P\Sigma P-\tau_\perp^T P\Sigma P\tau_\perp \tag{2.21}$$

where $\tau_\perp \triangleq I_n - \tau$. Note that in the full-order case $n_c = n$, $\tau = G = \Gamma = I_n$ and thus (2.18) and (2.19) reduce to the standard observer and regulator Riccati equations and (2.10)–(2.12) yield the usual LQG expressions. Furthermore, it can be shown that (2.20), (2.21), and (2.17) are equivalent to the assumption that (A_c, B_c, C_c) is controllable and observable.

Remark 2.3: Since $\hat{Q}\hat{P}$ is nonnegative semisimple it has a group generalized inverse $(\hat{Q}\hat{P})^\#$ given by $G^TM^{-1}\Gamma$ (see e.g., [21, p. 124]). Hence, by (2.9) the optimal projection τ is given by

$$\tau=\hat{Q}\hat{P}(\hat{Q}\hat{P})^\#. \tag{2.22}$$

Remark 2.4: The modified Riccati equations (2.13) and (2.14) are similar to the (single) "extended algebraic Riccati equation" which arises in the static output feedback problem (see, e.g., [22]).

Remark 2.5: Replacing x_c by Sx_c, where S is invertible, yields the "equivalent" compensator $(SA_cS^{-1}, SB_c, C_cS^{-1})$. Since $J(A_c, B_c, C_c) = J(SA_cS^{-1}, SB_c, C_cS^{-1})$ one would expect the Main Theorem to apply also to $(SA_cS^{-1}, SB_c, C_cS^{-1})$. This is indeed the case since transformation of the compensator state basis corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T\,(SMS^{-1})\,(S\Gamma)$. See [10] for related remarks.

Remark 2.6: By introducing the quasi-full-state estimate $\hat{x} \triangleq G^Tx_c \in \mathbb{R}^n$ so that $\tau\hat{x} = \hat{x}$ and $x_c = \Gamma\hat{x} \in \mathbb{R}^{n_c}$, (2.1)–(2.4) can be written as

$$\dot{x}=Ax+B\hat{C}_c\tau\hat{x}+w_1,$$

$$\dot{\hat{x}}=\tau(A-\hat{B}_cC+B\hat{C}_c)\tau\hat{x}+\tau\hat{B}_c(Cx+w_2)$$

where $\hat{B}_c \triangleq QC^TV_2^{-1}$ and $\hat{C}_c \triangleq -R_2^{-1}B^TP$. Although the *implemented* compensator has the state $x_c \in \mathbb{R}^{n_c}$, it can be viewed as a quasi-full-order compensator whose geometric structure is entirely dictated by the projection τ. Sensor inputs $\hat{B}_cy$ are annihilated unless they are contained in $[\mathfrak{N}(\tau)]^\perp = \mathfrak{R}(\tau^T)$, where $\mathfrak{N}$ and $\mathfrak{R}$ denote null space and range. Furthermore, the quasi-full-order state estimate $\tau\hat{x}$ employed in the control input is contained in $\mathfrak{R}(\tau)$. Thus, $\mathfrak{R}(\tau)$ and $\mathfrak{R}(\tau^T)$ are the control and observation subspaces of the compensator.

III. Proof of the Main Theorem

The proof given here considerably simplifies the original derivation given in [23] and [24]. Using the fact that $\mathfrak{a}_+$ is open, the Fritz John version of the Lagrange multiplier theorem can be used to rigorously derive the first-order necessary conditions ([7], see also [25])

$$0=\tilde{A}\tilde{Q}+\tilde{Q}\tilde{A}^T+\tilde{V}, \tag{3.1}$$

$$0=\tilde{A}^T\tilde{P}+\tilde{P}\tilde{A}+\tilde{R}, \tag{3.2}$$

$$0=P_{12}^TQ_{12}+P_2Q_2, \tag{3.3}$$

$$B_c=-(P_2^{-1}P_{12}^TQ_1+Q_{12}^T)C^TV_2^{-1}, \tag{3.4}$$

$$C_c=-R_2^{-1}B^T(P_1Q_{12}Q_2^{-1}+P_{12}), \tag{3.5}$$

where

$$\tilde{V} \triangleq \begin{bmatrix} V_1 & 0 \\ 0 & B_cV_2B_c^T \end{bmatrix}, \quad \tilde{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^TR_2C_c \end{bmatrix}$$

and $(n + n_c) \times (n + n_c)$ $\tilde{Q}$, $\tilde{P}$ are partitioned into $n \times n$, $n \times n_c$, and $n_c \times n_c$ subblocks as

$$\tilde{Q}=\begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \tilde{P}=\begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Expanding (3.1) and (3.2) yields

$$0=AQ_1+Q_1A^T+BC_cQ_{12}^T+Q_{12}(BC_c)^T+V_1, \tag{3.6}$$

$$0=AQ_{12}+Q_{12}A_c^T+BC_cQ_2+Q_1(B_cC)^T, \tag{3.7}$$

$$0=A_cQ_2+Q_2A_c^T+B_cCQ_{12}+Q_{12}^T(B_cC)^T+B_cV_2B_c^T, \tag{3.8}$$

$$0=A^TP_1+P_1A+(B_cC)^TP_{12}^T+P_{12}B_cC+R_1, \tag{3.9}$$

$$0=P_{12}A_c+A^TP_{12}+(B_cC)^TP_2+P_1BC_c, \tag{3.10}$$

$$0=A_c^TP_2+P_2A_c+(BC_c)^TP_{12}+P_{12}^TBC_c+C_c^TR_2C_c. \tag{3.11}$$

Writing (3.8) as (see [26], [27])

$$0=(A_c+B_cCQ_{12}Q_2^+)Q_2+Q_2(A_c+B_cCQ_{12}Q_2^+)^T+B_cV_2B_c^T$$

where Q_2^+ is the Moore–Penrose or Drazin generalized inverse of Q_2, it follows from [28, Lemmas 2.1 and 12.2] that Q_2 is positive definite. Similarly, (3.11) implies that P_2 is positive definite. This justifies (3.4) and (3.5).

Now define the $n \times n$ nonnegative-definite matrices (see [26], [27])

$$Q \triangleq Q_1-Q_{12}Q_2^{-1}Q_{12}^T, \qquad P \triangleq P_1-P_{12}P_2^{-1}P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12}Q_2^{-1}Q_{12}^T, \qquad \hat{P} \triangleq P_{12}P_2^{-1}P_{12}^T$$

and note that (3.3) implies (2.8) and (2.9) with

$$G \triangleq Q_2^{-1}Q_{12}^T,\ M \triangleq Q_2P_2,\ \Gamma \triangleq -P_2^{-1}P_{12}^T.$$

Since $Q_2P_2 = P_2^{-1/2}(P_2^{1/2}Q_2P_2^{1/2})P_2^{1/2}$, M is positive semisimple. Sylvester's inequality yields (2.17). Note also that

$$\hat{Q}=\tau\hat{Q},\ \hat{P}=\hat{P}\tau. \tag{3.12}$$

Next (2.11) and (2.12) follow from (3.4) and (3.5) by using the identities

$$Q_1=Q+\hat{Q},\ P_1=P+\hat{P}, \tag{3.13}$$

$$Q_{12}=\hat{Q}\Gamma^T,\ P_{12}=-\hat{P}G^T, \tag{3.14}$$

$$Q_2=\Gamma\hat{Q}\Gamma^T,\ P_2=G\hat{P}G^T. \tag{3.15}$$

Now substitute (2.11), (2.12), and (3.13)–(3.15) into (3.6)–(3.11) and use the relations

$$B_cC=\Gamma Q\Sigma,\ BC_c=-\Sigma PG^T,$$

$$B_cV_2B_c^T=\Gamma Q\Sigma Q\Gamma^T,\ C_c^TR_2C_c=GP\Sigma PG^T.$$

Then (2.10) follows from (3.8)–Γ(3.7). Substituting (2.10) into (3.7), (3.8), (3.10), and (3.11) shows that $((3.7)G)^T$ and $-(3.10)\Gamma$ are precisely (2.15) and (2.16). Since $G^T(3.8)G = (2.15)\tau$ and $\Gamma^T(3.11)\Gamma = \tau(2.16)$, (3.8) and (3.11) can be omitted. Finally, using (3.12) it follows that $(2.13) = (3.6) + (2.15)\tau - (2.15)–(2.15)^T$ and similarly for (2.14). □

IV. Directions for Further Research

With regard to the existence of a stabilizing compensator, known results (e.g., [28]–[34]) can be exploited to a great extent. A numerical algorithm for solving the optimal projection equations has been developed in [24] and [35]. The proposed computational scheme is philosophically quite different from gradient search algorithms [2], [3], [6], [7], [9], [11], [36], [37] in that it operates through direct solution of the optimal projection equations by iterative refinement of the optimal projection. Methods for eliminating local extrema are being investigated by applying component cost analysis [17]. Generalizations of the optimal projection equations can arise by considering the following extensions of the fixed-order dynamic-compensation problem.

1) *Discrete-Time System/Discrete-Time Compensator:* Digital implementation can be modeled by a discrete-time compensator with control of a continuous-time system facilitated by sampling and reconstruction devices.

2) *Cross Weighting/Correlated Disturbance and Observation Noise:* This extension is straighforward and entirely analogous to the LQG case (see, e.g., [3, p. 351]).

3) *Singular Observation Noise/Singular Control Weighting:* With due attention to (2.7), direct output feedback can be used in the singular case. The nature of the problem forebodes all of the difficulties associated with the singular LQG problem. Note that the output feedback problem [22], [38], when viewed in this context, is *highly* singular.

4) *Infinite-Dimensional Systems:* The optimal projection equations have been extended in [39] and [40] to the case in which (2.1) is a distributed parameter system, for example, a partial or functional differential equation.

5) *Decentralized Fixed-Order Controller:* The optimal projection equations can be derived for the case in which the dynamic controller has a fixed decentralized structure.

6) *Parameter Uncertainties:* The original derivation in [23] treated a Stratonovich state-dependent noise model representing parameter uncertainties in the plant. Further consideration of control- and measurement-dependent noise raises the possibility of directly including the impact of parameter uncertainties in the design of robust, implementable compensation for large-order systems.

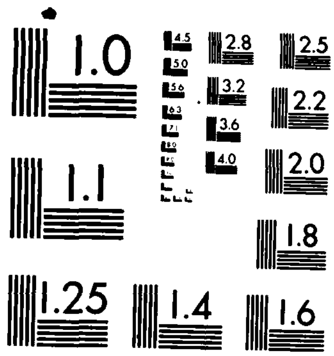
REFERENCES

[1] T. L. Johnson and M. Athans, "On the design of optimal constrained dynamic compensators for linear constant systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 658–660, 1970.
[2] W. S. Levine, T. L. Johnson, and M. Athans, "Optimal limited state variable feedback controllers for linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 785–793, 1971.
[3] K. Kwakernaak and R. Sivan, *Linear Optimal Control Systems.* New York: Wiley, 1972.
[4] D. B. Rom and P E. Sarachik, "The design of optimal compensators for linear constant systems wth inaccessible states," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 509–512, 1973.
[5] M. Sidar and B.-Z. Kurtaran, "Optimal low-order controllers for linear stochastic systems," *Int. J. Contr.*, vol. 22, pp. 377–387, 1975.
[6] J. M. Mendel and J. Feather, "On the design of optimal time-invariant compensators for linear stochastic time-invariant systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 653–657, 1975.
[7] S. Basuthakur and C. H. Knapp, "Optimal constant controllers for stochastic linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 664–666, 1975.
[8] R. B. Asher and J. C. Durrett, "Linear discrete stochastic control with a reduced-order dynamic compensator," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 626–627, 1976.
[9] W. J. Naeije and O. H. Bosgra, "The design of dynamic compensators for linear multivariable systems," in *Proc. 1977 IFAC,* Fredricton, N. B., Canada, 1977, pp. 205–212.
[10] P. J. Blanvillain and T. L. Johnson, "Invariants of optimal minimal-order observer based compensators," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 473–474, 1978.
[11] C. J. Wenk and C. H. Knapp, "Parameter optimization in linear systems with arbitrarily constrained controller structure," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 496–500, 1980.
[12] J. O'Reilly, "Optimal low-order feedback controllers for linear discrete-time systems," in *Control and Dynamic Systems,* Vol. 16, C. T. Leondes, Ed. New York: Academic, 1980.
[13] B. D. Moore, "Principal component analysis in linear systems: Controllability, observability and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17–32, 1981.
[14] E. I. Verriest, "Suboptimal LQG-design and balanced realizations," in *Proc. 20th IEEE Conf. Decision Contr.*, San Diego, CA, Dec. 1981, pp. 686–687.
[15] L. Pernebo and L. M. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382–387, 1982.
[16] E. A. Jonckheere and L. M. Silverman, "A new set of invariants for linear systems-application to reduced-order compensator design," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 953–964, 1983.
[17] A. Yousuff and R. E. Skelton, "Controller reduction by component cost analysis," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 520–530, 1984.
[18] D. C. Hyland and D. S. Bernstein, "The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore," in *23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984.
[19] D. A. Wilson, "Optimum solution of model-reduction problem," *Proc. IEE,* vol. 117, pp. 1161–1165, 1970.
[20] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications.* New York: Wiley, 1971.
[21] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations,* London: Pitman, 1979.
[22] J. Medanic, "On stabilization and optimization by output feedback," in *Proc. 12th Annual Asilomar Conf. Circuits and Syst.*, 1978, pp. 412–416.
[23] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," in *Proc. AIAA 20th Aerospace Sciences Meet.*, Orlando, FL, Jan. 1982, paper 82-0312.
[24] D. C. Hyland, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative result," in *AIAA 21st Aerospace Sciences Meet.*, Reno, NV, Jan. 1983, paper 83-0303.
[25] D. C. Hyland and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation," in *Proc. 22nd IEEE Conf. Decision Contr.*, San Antonio, TX, Dec. 1983.
[26] A. Albert, "Conditions for positive and nonnegative definiteness in terms of pseudo inverse," *SIAM J. Appl. Math.*, vol. 17, pp. 434–440, 1969.
[27] E. Kreindler and A. Jameson, "Conditions for nonnegativeness of partitioned matrices," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 147–148, 1972.
[28] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach.* New York: Springer-Verlag, 1974.
[29] F. M. Brasch and J. B. Pearson, "Pole placement using dynamic compensators," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 34–43, 1970.
[30] R. Ahmari and A. G. Vacroux, "On the pole assignment in linear systems with fixed-order compensators," *Int. J. Contr.*, vol. 17, pp. 397–404, 1973.
[31] D. C. Youla, J. J. Bongiorno, Jr., and C. N. Lu, "Single-loop feedback-stabilization of linear multivariable dynamical plants," *Automatica,* vol. 10, pp. 159–173, 1974.
[32] H. Seraji, "An approach to dynamic compensator design for pole assignment," *Int. J. Contr.*, vol. 21, pp. 955–966, 1975.
[33] R. V. Patel, "Design of dynamic compensators for pole assignment," *Int. J. Syst. Sci.*, vol. 7, pp. 207–224, 1976.
[34] C. A. Desoer, R. W. Liu, J. Murrary, and R. Saeks, "Feedback system design: The fractional representation approach to analysis and synthesis,' *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 399–412, 1980.
[35] D. C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 381–389.
[36] H. R. Sirisena and S. S. Choi, "Design of optimal constrained dynamic compensators for non-stationary linear stochastic systems," *Int. J. Contr.*, vol. 25, pp. 513–524, 1977.
[37] D. P. Looze and N. R. Sandell, Jr., "Gradient calculations for linear quadratic fixed control structure problems," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 285–288, 1980.
[38] W. S. Levine and M. Athans, "On the determination of the optimal constant output feedback gains for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 44–48, 1970.
[39] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of distributed parameter systems," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 396–400.
[40] D. S. Bernstein, "Explicit optimality conditions for finite-dimensional fixed-order dynamic compensation of infinite-dimensional systems," submitted for publication.

0018-9286/84/1100-1037$01.00

AD-A161 355 EXPLORATION OF THE MAXIMUM ENTROPY/OPTIMAL PROJECTION APPROACH TO CONTROL (U) HARRIS CORP MELBOURNE FL GOVERNMENT AEROSPACE SYSTEMS DIV D C HYLAND ET AL 2/3

UNCLASSIFIED FEB 85 AFOSR-TR-85-0966 F49620-84-C-0015 F/G 13/13 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

APPENDIX C

(REFERENCE [29])

THE OPTIMAL PROJECTION EQUATIONS FOR MODEL REDUCTION AND THE RELATIONSHIPS AMONG THE METHODS OF WILSON, SKELTON AND MOORE*

by

DAVID C. HYLAND and DENNIS S. BERNSTEIN

ABSTRACT

First-order necessary conditions for quadratically-optimal reduced-order modelling of linear time-invariant systems are derived in the form of a pair of modified Lyapunov equations coupled by an oblique projection which determines the optimal reduced-order model. This form of the necessary conditions considerably simplifies previous results of Wilson ([1]) and clearly demonstrates the quadratic extremality and nonoptimality of the balancing method of Moore ([2]). The possible existence of multiple solutions of the optimal projection equations is demonstrated and a relaxation-type algorithm is proposed for computing these local extrema. A cost component analysis of the model-error criterion similar to the approach of Skelton ([3]) is utilized at each iteration to direct the algorithm to the global minimum.

---

*This work was initially sponsored by the Department of the Air Force at Lincoln Laboratory/MIT, Lexington, MA, 02173-0073. Both authors are currently with Harris Corporation, Government Aerospace Systems Division, Controls Analysis and Synthesis Group, Melbourne, FL, 32901.

1. Introduction

The problem of approximating a high-order linear dynamical system with a relatively simpler system, i.e., the model-reduction problem, has received considerable attention in recent years. Among the myriad papers devoted to this problem are the notable contributions of Wilson ([1]), Moore ([2]) and Skelton ([3]) with which the present paper is concerned. In his 1970 paper, Wilson proposed an optimality-based approach to model reduction which involves minimizing the steady-state, quadratically-weighted output error when the original system and reduced-order model are subjected to white-noise inputs. For the resulting parameter optimization problem he obtained first-order necessary conditions which have the form of an aggregation (as, e.g., [4]) and which involve the solution of two unwieldly nonlinear matrix equations each of order $n+n_m$, where n and n_m are the orders of the original and reduced-order models, respectively ([5, 6]).

Some time later Moore proposed a quite different approach to model reduction based upon system-theoretic arguments as opposed to optimality criteria. Using the eigenvalues of the product of the controllability and observability gramians (which satisfy $n \times n$ Lyapunov equations), his method identifies subsystems which contribute little to the impulse response of the overall system. Such "weak" subsystems are thus eliminated to obtain a reduced-order model. This technique, known as balancing, has been vigorously developed in the recent literature ([7-11]). Since this approach is completely independent of any optimality considerations, however, there is no guarantee that such reduced-order models are in any sense optimal.

A third approach to model reduction, proposed by Skelton ([3,12]), also utilizes a quadratic optimality criterion as in [1]. However, rather than proceeding from necessary conditions as does Wilson, Skelton determines for a given basis the contribution (cost) of each state in a decomposition of the error criterion and truncates those with the least value. Although this approach is guided by optimality considerations, no rigorous guarantee of optimality is possible because of dependence on the choice of state space basis.

The present paper has five main objectives, the first of which is to show how the complex optimality conditions of Wilson (see (A.7)-(A.12) of [1]) can be transformed without loss of generality into much simpler and more tractable forms. The transformation is facilitated by exploiting the presence of an oblique (i.e., nonorthogonal) projection which was not recognized in [1] and which arises as a direct consequence of optimality. The resulting "optimal projection equations" constitute a coupled system of two nxn modified Lyapunov equations (see (2.13), (2.14) or (2.21), (2.22)) whose solutions are given by a pair of rank-n_m controllability and observability <u>pseudogramians</u>. The highly structured form of these equations gives crucial insight into the set of local extrema satisfying the first-order necessary conditions.

The second objective of the paper is to show how the optimal projection equations provide a rigorous extremality context for Moore's balancing method and to clearly demonstrate its nonoptimality. Although for some problems the "weak subsystem" hypothesis leads to a nearly optimal reduced-order model, we construct examples for which the reduced-order model obtained from the balancing method is much worse with respect to the least squares criterion than the quadratically-optimal reduced-order model. In general, all that can be said is that the presence of a weak subsystem indicates that the reduced-order model obtained by truncation in the balanced basis may be in the proximity of an extremal of the optimal model-reduction problem; however, this extremal may very well be a global maximum. It should be noted that in a recent paper ([13]), Kabamba has used bounds on the model error to demonstrate the quadratic nonoptimality of the balancing method.

The third objective of the paper is to demonstrate via an example the mechanism responsible for the existence of multiple extrema of the optimal model-reduction problem. By characterizing the optimal projection as a sum of rank-1 eigenprojections of the product of the rank-deficient pseudogramians, it is immediately clear that the first-order necessary conditions of the problem are ambiguous in the sense that they fail to specify <u>which</u> n_m eigenprojections comprise the optimal projection corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem. Specifically, since the pseudogramians can be rank deficient in $\binom{n}{n_m} = \frac{n!}{n_m!(n-n_m)!}$ ways, there may be precisely this many extremal projections corresponding to an identical number of local extrema.

The fourth objective of the paper is to propose a numerical algorithm for solving the optimal projection equations by exploiting their structure and taking advantage of the available insights. By expressing the modified Lyapunov equations in the form of "standard" Lyapunov equations, an iterative relaxation-type algorithm is developed. The crucial aspect of the proposed algorithm involves extracting an oblique projection at each step from the product of the Lyapunov equations. Since $\binom{n}{n_m}$ rank-n_m projections can be extracted from the product of two nxn positive-definite matrices, it is quickly evident that the criterion by which the n_m eigenprojections are chosen determines which of the numerous local extrema will be reached. If, for example, the projection is chosen in accordance with the n_m largest eigenvalues of the product of the solutions of the Lyapunov equations, then it should not be surprising in view of the previous discussion that a global maximum may very well be reached. In this case the first iteration of this algorithm involves Lyapunov equations whose solutions are the controllability and observability gramians and the eigenvalues in question are precisely the squares of the second-order modes ([2], p. 24). Thus the first iteration coincides with the (nonoptimal) balancing approach of [2].

Since the optimal projection equations are a consequence of differential (local) properties, it should not be expected that they alone would possess the inherent ability to identify the global minimum. Moreover, because of the number of local extrema, second-order necessary conditions appear to be useless. Instead, we investigate an approach which chooses the eigenprojections according to a cost-component analysis of the model-error criterion. This technique can lead to a global minimum by effectively eliminating the local extrema which have considerably greater cost than the global minimum. This approach is philosophically identical to the component cost analysis of Skelton ([3,12]). Essentially, then, component cost analysis is utilized at each iteration to direct the algorithm to the global minimum. Although our application of this technique is admittedly heuristic, it should be noted that it is essentially proposed as a device for efficiently "sorting out" the local extrema which satisfy the otherwise mathematically rigorous necessary conditions. Hence we propose component cost analysis as a crucial step in bridging the gap between local extremality and global optimality.

It should be pointed out that neither the numerical algorithm proposed in this paper nor the iterative algorithm developed in [4] and [5] has been proven to be convergent. The principal contribution of the present paper, however, is not a particular proposed algorithm but rather the revelations concerning the structure of the first-order necessary conditions. The proposed numerical algorithm should be considered but a prelude to a full investigation into numerical algorithms for the optimal projection equations. It should also be noted that the presence of the optimal projection was not exploited in developing the iterative algorithms in [4] and [5] (in fact, it was not even recognized in [1]) and hence crucial insight into local extrema was lacking.

The fifth and last objective of the paper is to point out the connection between the optimal projection equations for model reduction obtained herein and the first-order necessary conditions obtained recently for two closely related problems, namely, reduced-order state estimation and fixed-order dynamic compensation.

The plan of the paper is as follows. Section 2 begins with general notation and definitions followed by the model-reduction problem statement and the Main Theorem which presents the optimal projection equations for model reduction. A series of remarks considers various aspects of the Main Theorem and sets the stage for discussing connections with [1] and [2]. Section 3 contains a detailed discussion of the sense in which the optimal projection equations simplify the necessary conditions given in [1], and section 4 shows how the approach of [2] is approximately extremal. Section 5 presents a simple example which clearly displays the possible existence of multiple extrema satisfying the optimal projection equations. This example shows that the balancing method of [2] may lead to a nonoptimal reduced-order model and suggests a heuristic procedure for selecting the eigenprojections comprising the projection corresponding to the global minimum, i.e., the optimal projection. In section 6 a numerical algorithm for solving the optimal projection equations is proposed and applied to an illustrative example considered previously in [1] and [2] as well as to some interesting examples considered recently by Kabamba in [13]. Suggestions for further research are given in section 7 and the proof of the Main Theorem appears in the Appendix.

2. Problem Statement and Main Result

The following notation and definitions will be used throughout the paper:

I_r	rxr identity matrix
Z^T	transpose of vector or matrix Z
Z^{-T}	$(Z^T)^{-1}$ or $(Z^{-1})^T$
$\rho(Z)$	rank of matrix Z
tr Z	trace of square matrix Z
$\|\|Z\|\|$	$[\mathrm{tr}\, ZZ^T]^{1/2}$
Z_{ij}	(i,j)-element of matrix Z
$\mathrm{diag}(\alpha_1,\ldots,\alpha_r)$	rxr diagonal matrix with listed diagonal elements
E_i	matrix with unity in the (i,i) position and zeros elsewhere
$\underline{E}$	expected value
$\underline{R}$, $\underline{R}^{rxs}$	real numbers, rxs real matrices
stable matrix	matrix with eigenvalues in open left half plane
nonnegative-definite matrix	symmetric matrix with nonnegative eigenvalues
positive-definite matrix	symmetric matrix with positive eigenvalues
nonnegative-semisimple matrix	matrix similar to a nonnegative-definite matrix
positive-semisimple matrix	matrix similar to a positive-definite matrix
positive-diagonal matrix	diagonal matrix with positive diagonal elements
n, m, ℓ, n_m	positive integers, $1 \leq n_m \leq n$
x, u, y, x_m, y_m	n, m, ℓ, n_m, ℓ-dimensional vectors
A, B, C	nxn, nxm, ℓxn matrices
A_m, B_m, C_m	$n_m x n_m$, $n_m$xm, $\ell x n_m$ matrices
R, V	$\ell x \ell$, mxm positive-definite matrices

We consider the following problem.

Optimal Model-Reduction Problem. Given the controllable and observable system

$$\dot{x} = Ax + Bu, \tag{2.1}$$

$$y = Cx \tag{2.2}$$

find a reduced-order model

$$\dot{x}_m = A_m x_m + B_m u, \tag{2.3}$$

$$y_m = C_m x_m \tag{2.4}$$

which minimizes the quadratic model-reduction criterion*

$$J(A_m, B_m, C_m) \triangleq \lim_{t\to\infty} \underline{\underline{E}}[(y-y_m)^T R(y-y_m)],$$

where the input $u(t)$ is white noise with positive-definite intensity V. To guarantee that J is finite it is assumed that A is stable and we restrict our attention to the set of admissible reduced-order models

$$\underline{A} \triangleq \{(A_m, B_m, C_m): \ A_m \text{ is stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (2.3) and (2.4), we further restrict our attention to the set

$$\underline{A}_+ \triangleq \{(A_m, B_m, C_m) \in \underline{A}: \ (A_m, B_m) \text{ is controllable and } (A_m, C_m) \text{ is observable}\}.$$

*J will occasionally be referred to as the "model-reduction error" or, simply, as the "cost."

The following lemma is needed for the statement of the main result.

Lemma 2.1. Suppose $\hat{Q}, \hat{P} \in \underline{\underline{R}}^{n\times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_m$ then there exist $G, \Gamma \in \underline{\underline{R}}^{n_m \times n}$ and positive-semisimple $M \in \underline{\underline{R}}^{n_m \times n_m}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \tag{2.5}$$

$$\Gamma G^T = I_{n_m}. \tag{2.6}$$

Proof. By Theorem 6.2.5, p. 123 of [14] there exists nxn invertible $\hat{\Phi}$ such that the nonnegative-definite matrices $D_{\hat{Q}} \triangleq \hat{\Phi}\hat{Q}\hat{\Phi}^T$ and $D_{\hat{P}} \triangleq \hat{\Phi}^{-T}\hat{P}\hat{\Phi}^{-1}$ are both diagonal. Hence $D_{\hat{Q}}D_{\hat{P}}$ is nonnegative definite and $\hat{Q}\hat{P} = \hat{\Phi}^{-1}D_{\hat{Q}}D_{\hat{P}}\hat{\Phi}$ is nonnegative semisimple. Next introduce nxn orthogonal U to effect a rearrangement of basis if necessary so that

$$\hat{Q}\hat{P} = \Phi \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi^{-1},$$

where $\Phi \triangleq \hat{\Phi}U$ and $n_m \times n_m$ Λ is positive diagonal. Hence, for all $n_m \times n_m$ invertible S,

$$\hat{Q}\hat{P} = \Phi \begin{bmatrix} S \\ 0 \end{bmatrix} (S^{-1}\Lambda S)[S^{-1} \;\; 0]\Phi^{-1}$$

and thus (2.5) and (2.6) hold with $G = [S^T \;\; 0]\Phi^T$, $M = S^{-1}\Lambda S$ and $\Gamma = [S^{-1} \;\; 0]\Phi^{-1}$. □

For convenience in stating the Main Theorem we shall refer to $G, \Gamma \in \underline{\underline{R}}^{n_m \times n}$ and positive-semisimple $M \in \underline{\underline{R}}^{n_m \times n_m}$ satisfying (2.5) and (2.6) as a <u>(G,M,Γ)-factorization of $\hat{Q}\hat{P}$</u>. Also, define the positive-definite controllability and observability gramians

$$W_c \triangleq \int_0^\infty e^{At}BVB^Te^{A^Tt}dt,$$

$$W_o \triangleq \int_0^\infty e^{A^Tt}C^TRCe^{At}dt,$$

which satisfy the dual Lyapunov equations

$$0 = AW_c + W_cA^T + BVB^T, \tag{2.7}$$

$$0 = A^TW_o + W_oA + C^TRC. \tag{2.8}$$

Main Theorem. Suppose $(A_m, B_m, C_m) \in \underline{A_+}$ solves the optimal model-reduction problem. Then there exist nonnegative-definite matrices $\hat{Q}$, $\hat{P} \in \underline{\underline{R}}^{n\times n}$ such that, for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$, A_m, B_m and C_m are given by

$$A_m = \Gamma A G^T, \tag{2.9}$$

$$B_m = \Gamma B, \tag{2.10}$$

$$C_m = CG^T, \tag{2.11}$$

and such that, with $\tau \triangleq G^T\Gamma$, the following conditions are satisfied:

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_m, \tag{2.12}$$

$$0 = \tau[A\hat{Q} + \hat{Q}A^T + BVB^T], \tag{2.13}$$

$$0 = [A^T\hat{P} + \hat{P}A + C^TRC]\tau. \tag{2.14}$$

Several comments are in order. First note that the Main Theorem consists of necessary conditions in the form of two modified Lyapunov equations (2.13) and (2.14) plus rank conditions (2.12) which must possess nonnegative-definite solutions $\hat{Q}$, $\hat{P}$ when an optimal reduced-order model exists. We shall call $\hat{Q}$ and $\hat{P}$ the controllability and observability pseudogramians, respectively, since they are analogous to W_c and W_o and yet have rank deficiency. The modified Lyapunov equations are coupled by the nxn matrix τ which is a projection (idempotent matrix) since

$$\tau^2 = G^T\Gamma G^T\Gamma = G^T I_{n_m} \Gamma = \tau.$$

Note that in general $\mathcal{T}$ is an <u>oblique</u> projection and not necessarily an orthogonal projection since it may not be symmetric. We shall refer to a projection $\mathcal{T}$ corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem as an "optimal projection." It should be stressed that the form of the optimal reduced-order model (2.7)-(2.9) is a direct consequence of optimality and not the result of an a priori assumption on the structure of the reduced-order model.

Since the optimal projection equations are first-order necessary conditions for optimality, they may possess multiple solutions corresponding to various local extrema such as local maxima, local minima, saddle points, etc. The following definition will prove useful.

<u>Definition 2.1</u>. Nonnegative-definite $\hat{Q}$, $\hat{P} \in \underline{\underline{R}}^{n\times n}$ are <u>extremal</u> if (2.12)-(2.14) are satisfied. $(A_m, B_m, C_m) \in \underline{A}_+$ is <u>extremal</u> if there exist extremal $\hat{Q}$, $\hat{P}$ such that (A_m, B_m, C_m) is given by (2.9)-(2.11) for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$. The corresponding projection $\mathcal{T}$ is an <u>extremal projection</u>.

<u>Proposition 2.1</u>. Suppose (A_m, B_m, C_m) is extremal. Then the model-reduction error is given by

$$J(A_m, B_m, C_m) = 2\mathrm{tr}[(\hat{Q}\hat{P} - W_c W_o)A]. \tag{2.15}$$

<u>Proof</u>. The proof is given at the end of Appendix A. □

<u>Remark 2.1</u>. Noting the identities

$$-2\mathrm{tr}[W_c W_o A] = \mathrm{tr}[C^T R C W_c] = \mathrm{tr}[B V B^T W_o], \tag{2.16}$$

which follow from (2.7) and (2.8), (2.15) can be written for extremal (A_m, B_m, C_m) as

$$J(A_m, B_m, C_m) = 2\mathrm{tr}[\hat{Q}\hat{P}A] + \mathrm{tr}[C^T R C W_c] = 2\mathrm{tr}[\hat{Q}\hat{P}A] + \mathrm{tr}[B V B^T W_o]. \tag{2.17}$$

For convenience in the following discussion, let $\hat{Q}$, $\hat{P}$, G, M, Γ and τ correspond to some extremal (A_m, B_m, C_m). Now observe that if x_m is replaced by Sx_m then an "equivalent" reduced-order model is obtained with (A_m, B_m, C_m) replaced by $(SA_mS^{-1}, SB_m, C_mS^{-1})$. Since $J(A_m, B_m, C_m) = J(SA_mS^{-1}, SB_m, C_mS^{-1})$ one would expect the Main Theorem to apply also to $(SA_mS^{-1}, SB_m, C_mS^{-1})$. Indeed, the following result shows that this transformation corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T(SMS^{-1})(S\Gamma)$ and, moreover, that all (G, M, Γ)-factorizations of $\hat{Q}\hat{P}$ are related by a nonsingular transformation.

<u>Proposition 2.2</u>. If $S \in \underline{\underline{R}}^{n_m \times n_m}$ is invertible then $\bar{G} = S^{-T}G$, $\bar{\Gamma} = S\Gamma$ and $\bar{M} = SMS^{-1}$ satisfy

$$\hat{Q}\hat{P} = \bar{G}^T\bar{M}\bar{\Gamma}, \tag{2.5)'}$$

$$\bar{\Gamma}\bar{G}^T = I_{n_m}. \tag{2.6)'}$$

Conversely, if $\bar{G}, \bar{\Gamma} \in \underline{\underline{R}}^{n_m \times n}$ and invertible $\bar{M} \in \underline{\underline{R}}^{n_m \times n_m}$ satisfy (2.5)' and (2.6)' then there exists invertible $S \in \underline{\underline{R}}^{n_m \times n_m}$ such that $\bar{G} = S^{-T}G$, $\bar{\Gamma} = S\Gamma$ and $\bar{M} = SMS^{-1}$.

<u>Proof</u>. The first part is immediate. The second part follows by taking $S \triangleq \bar{M}^{-1}\bar{\Gamma}G^TM$, noting $S^{-1} = M\Gamma\bar{G}^T\bar{M}^{-1}$ and using the identities $\bar{\Gamma}G^TM\Gamma\bar{G}^T = \bar{M}$ and $M\Gamma\bar{G}^T = \bar{\Gamma}G^T\bar{M}$. □

The next result shows that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ.

<u>Proposition 2.3</u>. There exists invertible $\Phi \in \underline{\underline{R}}^{n \times n}$ such that

$$\hat{Q} = \Phi^{-1}\begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix}\Phi^{-T}, \qquad \hat{P} = \Phi^T\begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix}\Phi, \tag{2.18}$$

$$\hat{Q}\hat{P} = \Phi^{-1}\begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}\Phi, \qquad \tau = \Phi^{-1}\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}\Phi, \tag{2.19a,b}$$

where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \underline{\mathbb{R}}^{n_m \times n_m}$ are positive diagonal, $\Lambda \triangleq \Lambda_{\hat{Q}}\Lambda_{\hat{P}}$ and the diagonal elements of Λ are the eigenvalues of M. Consequently,

$$\hat{Q} = \tau\hat{Q}, \quad \hat{P} = \hat{P}\tau. \tag{2.20}$$

Proof. By Theorem 6.2.5, p. 123 of [14], and by (2.12) there exists nxn invertible Φ such that (2.18) holds and thus (2.19a) also holds. Define

$$\bar{G} = [I_{n_m} \quad 0]\Phi^{-T}, \quad \bar{M} = \Lambda \text{ and } \bar{\Gamma} = [I_{n_m} \quad 0]\Phi$$

so that (2.5)' and (2.6)' are satisfied. By the second part of Proposition 2.2 there exists invertible $S \in \underline{\mathbb{R}}^{n_m \times n_m}$ such that $G = S^T\bar{G}$, $M = S^{-1}\bar{M}S$ and $\Gamma = S^{-1}\bar{\Gamma}$. Now (2.19b) follows from

$$\tau = G^T\Gamma = \bar{G}^T\bar{\Gamma} = \Phi^{-1}\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}\Phi. \qquad \square$$

It is useful to present an alternative form of the optimal model-reduction equations (2.13) and (2.14). For convenience define the notation

$$\tau_{\perp} \triangleq I_n - \tau.$$

Proposition 2.4. Equations (2.13) and (2.14) are equivalent, respectively, to

$$0 = A\hat{Q} + \hat{Q}A^T + BVB^T - \tau_{\perp}BVB^T\tau_{\perp}^T, \tag{2.21}$$

$$0 = A^T\hat{P} + \hat{P}A + C^TRC - \tau_{\perp}^TC^TRC\tau_{\perp}. \tag{2.22}$$

Proof. By (2.20), (2.21) = (2.13) + $(2.13)^T$ + (2.13)τ and (2.13) = τ(2.21). Similarly, (2.14) and (2.22) are equivalent. $\square$

Remark 2.2. Noting the identities

$$-2\mathrm{tr}[\hat{Q}\hat{P}A] = \mathrm{tr}[C^TRC\hat{Q}] = \mathrm{tr}[BVB^T\hat{P}], \tag{2.23}$$

which follow from (2.20)-(2.22), (2.17) can be written for extremal (A_m, B_m, C_m) as

$$J(A_m, B_m, C_m) = \mathrm{tr}[C^T RC(W_c - \hat{Q})] = \mathrm{tr}[BVB^T(W_o - \hat{P}]. \tag{2.24}$$

To facilitate the discussion in the following sections, we consider the change of basis $\hat{x} \triangleq \Phi x$, where Φ is given by Proposition 2.3. Writing (2.1) and (2.2) as

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u, \tag{2.25}$$

$$y = \hat{C}\hat{x}, \tag{2.26}$$

where

$$\hat{A} \triangleq \Phi A \Phi^{-1},\ \hat{B} \triangleq \Phi B,\ \hat{C} \triangleq C\Phi^{-1},$$

(2.9) - (2.11) become

$$A_m = \hat{\Gamma}\hat{A}\hat{G}^T, \tag{2.27}$$

$$B_m = \hat{\Gamma}\hat{B}, \tag{2.28}$$

$$C_m = \hat{C}\hat{G}^T, \tag{2.29}$$

where

$$\hat{\Gamma} \triangleq \Gamma\Phi^{-1},\ \hat{G} \triangleq G\Phi^T$$

satisfy

$$\hat{G}^T\hat{\Gamma} = \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix},\ \hat{\Gamma}\hat{G}^T = I_{n_m}. \tag{2.30}$$

Note that (2.30) implies

$$\hat{\mathcal{T}} = [S \quad 0],\ \hat{G} = [S^{-T} \quad 0], \tag{2.31}$$

for some $n_m x n_m$ invertible S. Partitioning

$$\hat{x} = \begin{bmatrix} \hat{x}_m \\ \hat{x}_2 \end{bmatrix},\ \hat{A} = \begin{bmatrix} \hat{A}_m & \hat{A}_{m2} \\ \hat{A}_{2m} & \hat{A}_{22} \end{bmatrix},\ \hat{B} = \begin{bmatrix} \hat{B}_m \\ \hat{B}_2 \end{bmatrix},\ \hat{C} = [\hat{C}_m \quad \hat{C}_2],$$

where $\hat{x}_m \in \underline{\underline{R}}^{n_m}$ and $\hat{A}_m$, $\hat{B}_m$ and $\hat{C}_m$ are $n_m x n_m$, $n_m x m$ and $\ell x n_m$, respectively, (2.27)-(2.29) and (2.31) yield

$$A_m = S\hat{A}_m S^{-1},\ B_m = S\hat{B}_m,\ C_m = \hat{C}_m S^{-1}.$$

This shows that the optimal reduced-order model (modulo a state transformation) can be obtained by truncating the last $n-n_m$ states of the original system when it is expressed in the basis with respect to which $\hat{Q}$ and $\hat{P}$ have the diagonal forms $\begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix}$. Since the optimal projection $\mathcal{T}$ has the simple form $\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}$ in this basis, we shall refer to (2.25) and (2.26) as an <u>optimal projection realization</u> of (2.1) and (2.2). Note that when (2.21) and (2.22) are expanded in an optimal projection basis (i.e., a basis corresponding to an optimal projection realization) they assume the form

$$0 = \hat{A}_m \Lambda_Q + \Lambda_{\hat{Q}} \hat{A}_m^T + \hat{B}_m V \hat{B}_m^T, \tag{2.32}$$

$$0 = \hat{A}_{2m} \Lambda_{\hat{Q}} + \hat{B}_2 V \hat{B}_m^T, \tag{2.33}$$

$$0 = \hat{A}_m^T \Lambda_{\hat{P}} + \Lambda_{\hat{P}} \hat{A}_m + \hat{C}_m^T R \hat{C}_m, \tag{2.34}$$

$$0 = \Lambda_{\hat{P}} \hat{A}_{m2} + \hat{C}_m^T R \hat{C}_2. \tag{2.35}$$

If ϕ in Proposition 2.3 is replaced by

$$\begin{bmatrix} (\Lambda_{\hat{Q}} \Lambda_{\hat{P}})^{-1/4} & 0 \\ 0 & I_{n-n_m} \end{bmatrix} \phi,$$

which corresponds to a change of basis for the reduced-order model obtained by truncation, then $\Lambda_{\hat{Q}}$ and $\Lambda_{\hat{P}}$ are both replaced by $(\Lambda_{\hat{Q}}\Lambda_{\hat{P}})^{1/2}$ and hence this can be called a balanced optimal projection basis, utilizing the terminology of [2]. Thus, in a balanced optimal projection realization, $\Lambda_{\hat{Q}}$ and $\Lambda_{\hat{P}}$ appearing in (2.32)-(2.35) are equal.

The next result provides an interesting closed-form characterization of an extremal projection in terms of the Drazin generalized inverse of $\hat{Q}\hat{P}$. Since $(\hat{Q}\hat{P})^2 = G^T M^2 \Gamma$, and hence $\rho(\hat{Q}\hat{P})^2 = (\hat{Q}\hat{P})$, the "index" of $\hat{Q}\hat{P}$ (see [15], p. 121) is 1. In this case the Drazin inverse is traditionally called the group inverse and is denoted by $(\hat{Q}\hat{P})^{\#}$ ([15], p. 124). Since, as is easily verified, $(\hat{Q}\hat{P})^{\#} = G^T M^{-1}\Gamma$, (2.6) leads to the following result.

Proposition 2.5. An extremal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}. \tag{2.36}$$

An alternative representation for an extremal projection will prove useful for developing a numerical algorithm for solving (2.21) and (2.22). If $Q, P \in \underline{\underline{R}}^{r\times r}$ are nonnegative definite then by Lemma 2.1 QP is nonnegative semisimple and thus there exists invertible $\Psi \in \underline{\underline{R}}^{r\times r}$ such that

$$QP = \Psi^{-1}\Omega\Psi,$$

where $\Omega = \mathrm{diag}(\omega_1,\ldots,\omega_r)$ and $\omega_i \geq 0$ are the eigenvalues of QP. Now define the ith eigenprojection ([16], p. 41)

$$\Pi_i[QP] \triangleq \Psi^{-1}E_i\Psi,$$

which is a rank-1 oblique projection. Note that QP has the decomposition

$$QP = \sum_{i=1}^{r} \omega_i \Pi_i[QP].$$

Proposition 2.6. An extremal projection τ is given by

$$\tau = \sum_{i=1}^{n_m} \Pi_i[\hat{Q}\hat{P}], \tag{2.37}$$

where the ith eigenprojection $\Pi_i[\hat{Q}\hat{P}]$ corresponds to the ith nonzero eigenvalue λ_i of $\hat{Q}\hat{P}$.

3. Relationship to Wilson's Form of the Necessary Conditions

The optimal model-reduction problem considered in the previous section is identical to the problem considered by Wilson in [1] with the minor exception that he sets $R = I_\ell$. In [1] G and Γ are denoted by θ_2^T and θ_1, (2.6) appears as (15) and (2.9)-(2.11) are given by (14a,b). Note that in [1], θ_1 and θ_2 depend upon the solutions of a pair of $(n+n_m)\times(n+n_m)$ Lyapunov equations (see (7), (9) of [1] or (A.2), (A.3) of the present paper) whose coefficients and nonhomogeneous terms depend in turn on A_m, B_m and C_m (see (A.7)-(A.12)). The advantage of the nxn optimal projection equations (2.21) and (2.22) over the form of the necessary conditions given in [1] is that the former are independent of A_m, B_m and C_m. Moreover, the optimal projection τ, which was not recognized in [1], can be seen to play a fundamental role by coupling the modified Lyapunov equations (2.21) and (2.22) and determining (since $\tau = G^T\Gamma$) A_m, B_m and C_m in (2.7)-(2.9).

4. Relationship to Moore's Balancing Method

In contrast to Wilson's method for model reduction which is based on optimality principles, the approach due to Moore ([2]) relies on system-theoretic ideas. The main thrust of this approach "is to eliminate any weak subsystem which contributes little to the impulse response matrix" ([2], p. 26). The concept of a "weak subsystem" is defined by means of a dominance relation ([2], p. 28) involving similarity invariants called second-order modes. Moore evaluates reduced-order models obtained in this way by computing the relative error in the impulse response given for MIMO systems by ([2], p. 29)

$$\epsilon(A_m,B_m,C_m) \triangleq \left[\int_0^\infty ||H_e(t)||^2 dt \Big/ \int_0^\infty ||H(t)||^2 dt\right]^{1/2},$$

where $H_e(t) \triangleq H(t)-H_m(t)$, $H(t) \triangleq R^{1/2}Ce^{At}BV^{1/2}$ and $H_m(t) \triangleq R^{1/2}C_m e^{A_m t}B_m V^{1/2}$.
To discuss this approach in the context of the optimal model-reduction problem we assume that $V = I_m$ and $R = I_\ell$.

Proposition 4.1. Suppose $(A_m, B_m, C_m) \in \underline{A}$. Then

$$\begin{aligned}\epsilon(A_m, B_m, C_m) &= [\tfrac{1}{2} J(A_m, B_m, C_m)/\mathrm{tr}(W_c W_o A]^{1/2} \\ &= [J(A_m, B_m, C_m)/\mathrm{tr}(C^T R C W_c)]^{1/2} \\ &= [J(A_m, B_m, C_m)/\mathrm{tr}(B V B^T W_o)]^{1/2}.\end{aligned} \tag{4.1}$$

Proof. The result follows from (A.1), (A.8) and (A.9) which hold without regard to either optimality or extremality.□

Note that Proposition 4.1 shows that the relative error in the impulse response is minimized precisely when $J(A_m, B_m, C_m)$ is minimized. Actually, this result is to be expected since, as shown in [1], J can be obtained alternatively by taking u(t) to be an impulse at t = 0.

To draw interesting comparisons with the results of [2], choose nxn invertible Ψ such that $\Psi W_c \Psi^T$ and $\Psi^{-T} W_o \Psi^{-1}$ are both diagonal and hence

$$W_c W_o = \Psi^{-1} \Sigma^2 \Psi, \tag{4.2}$$

where $\Sigma \triangleq \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ and the second-order modes σ_i (i.e., the positive square roots of the eigenvalues of $W_c W_o$) satisfy $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$. This transformation corresponds to replacing (2.1), (2.2) by

$$\dot{\overline{x}} = \overline{A}\overline{x} + \overline{B}u, \tag{4.3}$$

$$y = \overline{C}\overline{x}, \tag{4.4}$$

where

$$\overline{x} \triangleq \Psi x, \ \overline{A} \triangleq \Psi A \Psi^{-1}, \ B \triangleq \Psi B, \ \overline{C} \triangleq C \Psi^{-1}. \tag{4.5}$$

The transformed system (4.3), (4.4), called a principal axis realization ([17]), can further be chosen so that

$$\Psi W_c \Psi^T = \Psi^{-T} W_o \Psi^{-1} = \Sigma, \tag{4.6}$$

i.e., the balanced realization. Using (4.5), (2.7) and (2.8) become

$$0 = \bar{A}\Sigma + \Sigma\bar{A}^T + \bar{B}V\bar{B}^T, \tag{4.7}$$

$$0 = \bar{A}^T\Sigma + \Sigma\bar{A} + \bar{C}^TR\bar{C}. \tag{4.8}$$

The model-reduction procedure suggested in [2] involves partitioning

$$\bar{x} = \begin{bmatrix} \bar{x}_m \\ \bar{x}_2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} \bar{A}_m & \bar{A}_{m2} \\ \bar{A}_{2m} & \bar{A}_{22} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \bar{B}_m \\ \bar{B}_2 \end{bmatrix}, \quad \bar{C} = [\bar{C}_m \quad \bar{C}_2],$$

where $\bar{x}_m \in \underline{\underline{R}}^{n_m}$ and A_m, B_m and C_m have corresponding dimension, and extracting the reduced-order model (A_m, B_m, C_m). Hence the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ is extracted from (4.3), (4.4) in essentially the same way the optimal reduced-order model (A_m, B_m, C_m) is extracted from (2.25), (2.26). To see how the optimal-projection realization compares to a principal-axis realization, first note that (2.13) and (2.14) are satisfied by $\hat{Q} = W_c$ and $\hat{P} = W_o$ when the rank conditions (2.10) are ignored. Indeed, since W_c and W_o are positive definite, the rank conditions (2.12) do <u>not</u> hold. If, however, the system (2.1), (2.2) is expressed in the balanced coordinate system (4.3), (4.4) (so that $W_c = W_o = \Sigma$), then the assumption $\sigma_{n_m} \gg \sigma_{n_m+1}$ implies that $\rho(W_c)$, $\rho(W_o)$ and $\rho(W_cW_o)$ are "approximately" equal to n_m and thus, in this sense, condition (2.10) <u>is</u> satisfied. This observation leads to the suggestion that when $\sigma_{n_m} \gg \sigma_{n_m+1}$, W_c and W_o are approximations to solutions $\hat{Q}$ and $\hat{P}$ of the optimal projection equations and the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ of Moore is an approximation to some extremal (A_m, B_m, C_m). There is no guarantee, of course, that any particular extremum corresponds to the global minimum, or even to a local minimum.

5. <u>Existence of Multiple Extrema and Cost-Component Ranking</u>

In this section we show by means of a simple example that the optimal projection equations may possess nonunique solutions corresponding to multiple extrema, e.g., local minima or maxima. We also show how decomposing the cost can identify the global minimum from amongst the numerous extrema. To begin, let $m = \ell = n$, $R = V = I_n$,

$$A \triangleq \mathrm{diag}(-\alpha,\ldots,-\alpha_n),$$

where $\alpha_i > 0$, $i=1,\ldots,n$, and suppose B and C are such that

$$BB^T = \mathrm{diag}(\beta_1,\ldots,\beta_n),\quad C^TC = \mathrm{diag}(\gamma_1,\ldots,\gamma_n),$$

where $\beta_i > 0$, $\gamma_i > 0$, $i=1,\ldots,n$. Hypothesizing diagonal solutions $\hat{Q}$ and $\hat{P}$ of (2.21) and (2.22) leads to

$$\hat{Q}_{ii} = \frac{\beta_i}{2\alpha_i}\delta_i,\quad \hat{P}_{ii} = \frac{\gamma_i}{2\alpha_i}\delta_i,$$

where each δ_i, $i=1,\ldots,n$ is either zero or one and exactly n_m of the δ_i's are equal to one. Hence $\tau = \mathrm{diag}(\delta_1,\ldots,\delta_n)$. Note that there are $\binom{n}{n_m}$ such solutions of the optimal projection equations corresponding to $\binom{n}{n_m}$ local extrema.

Since $W_c = -\frac{1}{2}A^{-1}BB^T$, $W_o = -\frac{1}{2}A^{-1}C^TC$, $\hat{Q} = \tau W_c$, $\hat{P} = \tau W_o$

and A, W_c and W_o commute, (2.15) becomes

$$J(A_m,B_m,C_m) = -\frac{1}{2}\,\mathrm{tr}\,\tau_\perp A^{-1}BB^TC^TC.$$

Hence

$$J(A_m,B_m,C_m) = \sum_{i=1}^{n} \zeta_i(1-\delta_i), \tag{5.1}$$

where

$$\zeta_i \triangleq \beta_i\gamma_i/2\alpha_i.$$

To minimize J it is clear that δ_i should be chosen to be unity for the largest n_m elements of the set $\left\{\zeta_i\right\}_{i=1}^{n}$ and zero otherwise. Although this choice is not necessarily unique, it does yield a global minimum. Note that choosing $\delta_i = 1$ is equivalent to selecting a particular eigenprojection $\Pi_i[W_c W_o]$ corresponding to the eigenvalue $\beta_i \gamma_i / 4\alpha_i^2$.

Remark 5.1. The expression in (5.1) can be regarded as a decomposition of the cost in terms of the state variables. The idea of deleting states based on their "component costs" is precisely the "component cost analysis" approach of Skelton ([3,12]).

Using the example it is easy to see that the balancing method of [2], which selects eigenprojections based upon the magnitude of the eigenvalues of $W_c W_o$, i.e., the (squares of the) second-order modes, may yield a grossly suboptimal reduced-order model. To this end let

$$\alpha_1 = 1, \ \alpha_2 = 10^6, \ \beta_1 = 1, \ \beta_2 = 10^6, \ \gamma_1 = 1, \ \gamma_2 = 10^3$$

so that

$$\zeta_1 = .5, \qquad \zeta_2 = 500.$$

Clearly J is minimized ($J = \zeta_1$) by choosing $\delta_1 = 0$, $\delta_2 = 1$, which corresponds to truncating the first state variable. If, however, the method of [2] is utilized, then judging by the second-order modes

$$\sigma_1 = .5, \qquad \sigma_2 = (2.5)^{1/2} \cdot 10^{-2} \approx .012,$$

the second state variable should be deleted. This, however, corresponds to choosing $\delta_1 = 1$, $\delta_2 = 0$ with the much higher cost $J = \zeta_2$. The fact that the balancing approach of [2] fails to determine a solution of the optimal model-reduction problem should not be surprising in view of the fact that the error criterion plays no role in the balancing technique.

Although the above solution exploited the simple structure of this example, it is clear that choosing the global minimum from amongst the local extrema involves an eigenprojection decomposition of the cost J. To extend this idea to more general systems, we invoke the following heuristic approximation.

Approximation 5.1. Let Ψ define the balanced basis as in (4.6). Then Ψ also approximately defines a balanced optimal projection basis, i.e.,

$$\Psi\hat{Q}\Psi^T \approx \Psi^{-T}\hat{P}\Psi^{-1} \approx \bar{\tau}\Sigma^2, \tag{5.2}$$

where extremal

$$\bar{\tau} \triangleq \Psi\tau\Psi^{-1} = \mathrm{diag}(\delta_1,\ldots,\delta_n) \tag{5.3}$$

and

$$\delta_i \in \{0,1\}, \qquad \sum_{i=1}^{n} \delta_i = n_m.$$

Proposition 5.1. If approximation 5.1 holds for extremal (A_m,B_m,C_m) then, with $\bar{\tau}_\perp \triangleq I_n - \bar{\tau}$,

$$\begin{aligned} J(A_m,B_m,C_m) &\approx -2\mathrm{tr}[\bar{\tau}_\perp\Sigma^2\bar{A}] && (5.4)\\ &= 2\sum_{i=1}^{n} -\sigma_i^2\bar{A}_{ii}(1-\delta_i). \end{aligned}$$

Remark 5.2. From (4.7) and (4.8) it follows that (5.4) can be written either as

$$\begin{aligned} J(A_m,B_m,C_m) &\approx \mathrm{tr}[\bar{\tau}_\perp\Sigma\bar{B}V\bar{B}^T] && (5.5)\\ &= \sum_{i=1}^{n} \sigma_i(\bar{B}V\bar{B}^T)_{ii}(1-\delta_i) \end{aligned}$$

or

$$J(A_m, B_m, C_m) \approx \mathrm{tr}[\bar{\tau}_{\perp}\Sigma\bar{C}^T R\bar{C}] \qquad (5.6)$$

$$= \sum_{i=1}^{n} \sigma_i (\bar{C}^T R\bar{C})_{ii}(1-\delta_i).$$

Hence, Approximation 5.1 leads to the following cost-component ranking (again, in the sense of Skelton [3,12]) of the $\binom{n}{n_m}$ extrema satisfying the optimal projection equations.

Cost-Component Ranking. Assume Approximation 5.1 is valid and choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$\delta_i = 1$, if $-\sigma_i^2\bar{A}_{ii}$ is among the n_m largest elements of the set $\left\{-\sigma_r^2\bar{A}_{rr}\right\}_{r=1}^{n}$;

$\delta_i = 0$, otherwise.

For comparison purposes we shall also consider the following ranking of the eigenprojections based upon the eigenvalues of $W_c W_o$ (i.e., second-order modes).

Eigenvalue Ranking. Choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$\delta_i = 1$, if σ_i is among the n_m largest elements of the set $\left\{\sigma_r\right\}_{r=1}^{n}$;

$\delta_i = 0$, otherwise.

Remark 5.3. The observation that the second-order modes alone may be a poor guide to determining an optimal reduced-order model has recently been made in [13] where bounds on the model-error criterion were given involving both the second-order modes and suitable weights called balanced gains. It can be seen from Proposition 5.1 that the role of balanced gains in our approach is played by the elements $-\sigma_i\bar{A}_{ii}$ when Approximation 5.1 holds. It can also be seen that the balanced gains of Kabamba yield bounds on the component costs of Skelton.

6. Numerical Solution of the Optimal Projection Equation

Insofar as the ultimate aim of any model-reduction technique is to permit the development of numerical procedures for reducing high-order models, the optimal projection equations, comprising a coupled system of modified Lyapunov equations, appear promising in this regard. Therefore, we present an iterative computational algorithm that exploits the structure of these equations and the available insights. The reader is strongly reminded that the proposed algorithm is but a first attempt at solving these new equations and alternative algorithms may yet be devised. The basis of this algorithm is the ability to write the modified Lyapunov equations (2.21), (2.22) in the form of "standard" Lyapunov equations (6.1), (6.2) such that the pseudogramians $\hat{Q}$ and $\hat{P}$ are extracted at the final step (6.6). It follows from (2.32)-(2.35) that (2.21), (2.22) are indeed equivalent to (6.1), (6.2) (with $k = \infty$) and (6.6).

Algorithm:

Step 1: Initialize $\mathcal{T}^{(0)} = I_n$;

Step 2: Solve for $\hat{Q}^{(k)}$, $\hat{P}^{(k)}$:

$$0 = (A - \mathcal{T}^{(k)}A\mathcal{T}_1^{(k)})\hat{Q}^{(k)} + \hat{Q}^{(k)}(A - \mathcal{T}^{(k)}A\mathcal{T}_1^{(k)})^T + BVB^T, \tag{6.1}$$

$$0 = (A - \mathcal{T}_1^{(k)}A\mathcal{T}^{(k)})^T\hat{P}^{(k)} + \hat{P}^{(k)}(A - \mathcal{T}_1^{(k)}A\mathcal{T}^{(k)}) + C^TRC; \tag{6.2}$$

Step 3: Balance:

$$\Phi^{(k)}\hat{Q}^{(k)}(\Phi^{(k)})^T = (\Phi^{(k)})^{-T}\hat{P}^{(k)}(\Phi^{(k)})^{-1} = \Sigma^{(k)}, \tag{6.3}$$

$$\Sigma^{(k)} = \mathrm{diag}(\sigma_1^{(k)},\dots,\sigma_n^{(k)}),\quad \sigma_1^{(k)} \geq \sigma_2^{(k)} \geq \dots \geq \sigma_n^{(k)} \geq 0;$$

Step 4: If $k > 1$ check for convergence:

$$e_k \triangleq \left[\frac{\mathrm{tr}(C^TRCW_c) - \mathrm{tr}(C^TRC\mathcal{T}^{(k)}\hat{Q}^{(k)}(\mathcal{T}^{(k)})^T)}{\mathrm{tr}(C^TRCW_c)}\right]^{\frac{1}{2}} \tag{6.4}$$

If $|e_k - e_{k-1}| <$ tolerance then go to Step 8;

Else continue;

Step 5: Select n_m eigenprojections:

$$\Pi_{i_1}^{(k)}[\hat{Q}^{(k)}\hat{P}^{(k)}], \ldots, \Pi_{i_{n_m}}^{(k)}[\hat{Q}^{(k)}\hat{P}^{(k)}],$$

$$\Pi_i[\hat{Q}^{(k)}\hat{P}^{(k)}] \triangleq S^{(k)}E_i(S^{(k)})^{-1};$$

Step 6: Update:

$$\mathcal{T}^{(k+1)} = \sum_{r=1}^{n_m} \Pi_{i_r}^{(k)}[\hat{Q}^{(k)}\hat{P}^{(k)}]; \tag{6.5}$$

Step 7: Check for convergence; if not, increment k and return to Step 2;

Step 8: Set:

$$\hat{Q} = \mathcal{T}^{(\infty)}\hat{Q}(\mathcal{T}^{(\infty)})^T, \ \hat{P} = (\mathcal{T}^{(\infty)})^T\hat{P}\mathcal{T}^{(\infty)}. \tag{6.6}$$

For convenience we shall adopt the notation $(A_m^{(k)}, B_m^{(k)}, C_m^{(k)})$, where $k > 0$, to denote the reduced-order model obtained as a result of applying the projection $\mathcal{T}^{(k)}$ and we define

$$\epsilon_k \triangleq \epsilon(A_m^{(k)}, B_m^{(k)}, C_m^{(k)}),$$

i.e., the relative error associated with $(A_m^{(k)}, B_m^{(k)}, C_m^{(k)})$. Note that in general $\epsilon_k \neq e_k$ since e_k denotes the relative error only when convergence has been reached.

It should be clear from the discussion in the previous section that the crucial step of the algorithm is Step 5, the choice of the eigenprojections. For

the examples which follow we shall invoke consistently at Step 5 either the cost-component ranking based upon Approximation 5.1 or the eigenvalue ranking.

Remark 6.1. Note that in the special case $R = I_m$ and $V = I_\ell$, the first iteration of the algorithm yields $\hat{Q}^{(0)} = W_c$, $\hat{P}^{(0)} = W_o$. If, at Step 5, we choose $i_r = r$, $r = 1,\ldots,n_m$, i.e., the eigenprojections are selected according to the eigenvalue ranking, then $(A_m^{(1)}, B_m^{(1)}, C_m^{(1)})$ is precisely the reduced-order model obtained from balancing.

We shall first consider the following example which was treated by both Wilson and Moore. In this example and those that follow assume $R = I_m$, $V = I_\ell$.

Example 6.1

$$A = \begin{bmatrix} 0 & 0 & 0 & -150 \\ 1 & 0 & 0 & -245 \\ 0 & 1 & 0 & -113 \\ 0 & 0 & 1 & -19 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad C = [0 \quad 0 \quad 0 \quad 1].$$

Table 1 summarizes the results obtained for the three cases $n_m = 3,2,1$ utilizing the eigenvalue ranking. In each case the proposed algorithm converged linearly in less than eight iterations and in each case improvement is evident over previously published results. As pointed out in [2], Wilson's result seems to imply a lack of final convergence. For this example the balancing approach yields a reduced-order model close to the global minimum.

Table 1. Relative Error $e_\infty = \epsilon_\infty$

Order n_m	Wilson [1]	Moore [2]	Optimal Projection Equations
3	-	.001311	.001306
2	.04097	.03938	.03929
1	-	.4321	.04268

We now turn to a pair of interesting examples considered in [13].

Example 6.2.

$$A = \begin{bmatrix} -.005 & -.99 \\ -.99 & -5000 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 100 \end{bmatrix}, \quad C = B^T.$$

Table 2 summarizes the results obtained using the eigenvalue ranking and Table 3 gives the results when the cost-component ranking is used. It is clear that the former method directs the algorithm to the global maximum whereas the latter approach yields the global minimum.

Example 6.3.

$$A = \begin{bmatrix} -.25 & -.4 \\ -.4 & -.72 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1.2 \end{bmatrix}, \quad C = B^T.$$

Table 4 reports the results obtained using either the cost-component ranking or the eigenvalue ranking which agree for this example. If the alternative eigenprojection is selected then, as expected, the algorithm converges to a global maximum (see Table 5). The interesting aspect of this example, as discussed in [13], is that the error $\epsilon_1 = .5245$ for the reduced-order model obtained by either eigenprojection ranking is actually greater than $\epsilon_1 = .3849$ obtained by choosing the alternative reduced-order model. This situation seems to indicate that proper eigenprojection selection based upon a cost decomposition is able to direct the algorithm to the global minimum in cases for which the starting values are not nearby.

Table 2. Example 6.2 with Eigenvalue Ranking

k	e_k
1	.9950371897
2	.9950371691
3	.9950371690

Table 3. Example 6.2 with Cost-Component Ranking

k	e_k
1	.0995037
2	.0995449
3	.0995924
4	.0996520
5	.0997346
6	.0998648
7	.1001125
8	.1007724
9	.1054569
10	.0982006
11	.0975409
12	.0975342
13	.0975330
14	.0975329

Table 4. Example 6.3 Using Either Ranking

k	e_k
1	.646996
2	.418341
3	.220994
4	.177276
5	.176576

Table 5. Example 6.3 with the Opposite Ranking

k	e_k
1	.7624928516
2	.9999999961
3	.9999999975
.	.
.	.
.	.
29	.9999999999

7. The Optimal Projection Equations for Fixed-Order Dynamic Compensation and Reduced-Order State Estimation

We briefly discuss the relationship between the optimal projection equations for model reduction and analogous results for reduced-order control and estimation problems.

Fixed-Order Dynamic-Compensation Problem. Given the control system

$$\dot{x} = Ax + Bu + w_1, \tag{7.1}$$
$$y = Cx + w_2 \tag{7.2}$$

design a fixed-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \tag{7.3}$$
$$u = C_c x_c \tag{7.4}$$

which minimizes the performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \underline{\underline{E}}[x^T R_1 x + u^T R_2 u], \tag{7.5}$$

where $u \in \underline{\underline{R}}^m$, $x_c \in \underline{\underline{R}}^{n_c}$, $n_c \leq n$, w_1 is white disturbance noise, w_2 is nonsingular white observation noise, R_1 is nonnegative definite and R_2 is positive definite.

Necessary conditions characterizing optimal (A_c, B_c, C_c) have been developed in [18-22] along the same lines as the Main Theorem. These conditions, called the optimal projection equations for fixed-order dynamic compensation, consist of four matrix equations (two modified Riccati equations and two modified Lyapunov equations) coupled by a projection. The modified Riccati equations, not surprisingly, are similar in form to the covariance and cost Riccati equations of LQG theory and the modified Lyapunov equations are similar to the optimal model-reduction equations (2.13) and (2.14). Hence, while the modified Riccati equations govern optimal estimation and optimal control, the additional modified Lyapunov equations characterize "optimal reduction". The important fact that all four equations are coupled supports the view that optimal fixed-order dynamic compensators <u>cannot</u> in general be designed by means of a stepwise procedure, e.g., by either open-loop model reduction followed by LQG or LQG followed by closed-loop model reduction.

Mid way between the model-reduction and fixed-order dynamic-compensation problems lies the following problem.

<u>Reduced-Order State-Estimation Problem</u>. Given the system

$$\dot{x} = Ax + w_1, \tag{7.6}$$
$$y = Cx + w_2, \tag{7.7}$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \tag{7.8}$$
$$y_e = C_e x_e, \tag{7.9}$$

which minimizes the estimation criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t\to\infty} \underline{\underline{E}}\,[(Lx-y_e)^T R(Lx-y_e)],$$

where $x_e \in \underline{\underline{R}}^{n_e}$, $L \in \underline{\underline{R}}^{p\times n_e}$ and L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (7.8) and (7.9) in real time.

In view of the results already given it should not be surprising (see [23]) that the optimal projection equations for reduced-order state estimation form a system of <u>three</u> matrix equations (a pair of modified Lyapunov equations along with a single modified Riccati equation) coupled by a projection which determines the gains of the optimal reduced-order estimator. This intrinsic coupling between the "operations" of optimal estimation (the modified Riccati equation) and optimal model reduction (the pair of modified Lyapunov equations) stresses the fact that reduced-order estimators designed by means of either model reduction followed by "full-order" state estimation or full-order estimation followed by estimator reduction will generally not be optimal for the given order.

8. Directions for Further Research

The most important area of research involves the further development of algorithms for solving the optimal projection equations. Although proving local convergence of the proposed algorithm appears possible, the more important problem is achieving global optimality via the component cost approach. Although the global minimum was attained for all examples attempted by the authors, it remains to treat considerably more complex systems.

An interesting extension of the Main Theorem involves the case in which the original system (2.1), (2.2) is a distributed parameter system, e.g., a partial differential equation or a functional differential equation. This generalization, which has been referred to as the "ultimate reduced-order problem" ([24]), may lead to the efficient generation of high-order discretizations for such systems. All of the mathematical machinery required to generalize the Main Theorem to this case has already been applied to fixed-order dynamic compensation in ([25]).

9. Conclusion

First-order necessary conditions for quadratically optimal reduced-order modelling of a linear time-invariant plant are expressed in the form of a pair of nxn modified Lyapunov equations coupled by an oblique projection. This form of the necessary conditions considerably simplifies the original form given by Wilson in [1] and clearly reveals the possible presence of numerous extrema. The balancing method of Moore given in [2] is shown to yield a reduced-order model that is "close" to an extremal given by the necessary

conditions. A numerical example shows, however, that this extremal may very well be the global maximum rather than the desired global minimum. An algorithm is proposed which exploits the presence of the optimal projection and computes the various local extrema by the choice of eigenprojections comprising the projection. A component-cost ranking of the eigenprojections, which is very much in the spirit of Skelton's method in [3,12], is used to direct the algorithm to the global optimum.

Appendix: Proof of the Main Theorem

Introducing the augmented system

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u,$$
$$\tilde{y} = \tilde{C}\tilde{x},$$

where

$$\tilde{x} \triangleq \begin{bmatrix} x \\ x_m \end{bmatrix}, \quad \tilde{y} \triangleq y - y_m,$$

$$\tilde{A} \triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \quad \tilde{B} \triangleq \begin{bmatrix} B \\ B_m \end{bmatrix}, \quad \tilde{C} \triangleq [C \quad -C_m],$$

leads to the expression

$$J(A_m, B_m, C_m) = \mathrm{tr}\, \tilde{Q}\tilde{R}, \tag{A.1}$$

where $\tilde{R} \triangleq \tilde{C}^T R \tilde{C}$ and the nonnegative-definite steady-state covariance $\tilde{Q}$ of $\tilde{x}$ is given by the (unique) solution of

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}, \tag{A.2}$$

with $\tilde{V} \triangleq \tilde{B}V\tilde{B}^T$. To minimize (A.1) subject to the constraint (A.2), form the Lagrangian

$$L(A_m, B_m, C_m, \tilde{Q}) \triangleq \mathrm{tr}[\lambda\tilde{Q}\tilde{R} + (\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V})\tilde{P}]$$

with multipliers $\lambda \geq 0$ and $\widetilde{P} \in \underline{\underline{R}}^{(n+n_m)x(n+n_m)}$. Since $\underline{A}_+$ is an open set the standard Lagrange multiplier rule can be applied.

Using formulas for computing partial derivatives ([26]) it follows that

$$0 = L_{\widetilde{Q}} = \widetilde{A}^T\widetilde{P} + \widetilde{P}\widetilde{A} + \lambda\widetilde{R}.$$

Since $\lambda = 0$ implies $\widetilde{P} = 0$ (recall $\widetilde{A}$ is stable), we can take $\lambda = 1$ without loss of generality. Hence $\widetilde{P}$ is the (unique nonnegative-definite) solution of

$$0 = \widetilde{A}^T\widetilde{P} + \widetilde{P}\widetilde{A} + \widetilde{R}. \tag{A.3}$$

Again using formulas from [26] and performing some manipulation it follows that

$$0 = L_{A_m} = Q_{12}^T P_{12} + Q_2 P_2, \tag{A.4}$$

$$0 = L_{B_m} = 2(P_{12}^T B + P_2 B_m)V, \tag{A.5}$$

$$0 = L_{C_m} = 2R(C_m Q_2 - CQ_{12}), \tag{A.6}$$

where $\widetilde{Q}$ and $\widetilde{P}$ have been partitioned as

$$\widetilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \widetilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}. \tag{A.7}$$

Since (as will be seen shortly) Q_2 and P_2 are positive definite, define

$$G \triangleq Q_2^{-1}Q_{12}^T, \quad \Gamma \triangleq -P_2^{-1}P_{12}^T, \tag{A.8}$$

so that (A.4)-(A.6) become (2.6), (2.10) and (2.11), respectively.

Next, define the nonnegative-definite matrices

$$\hat{Q} \triangleq Q_{12}Q_2^{-1}Q_{12}^T, \quad \hat{P} \triangleq P_{12}P_2^{-1}P_{12}^T \tag{A.9}$$

and note that (A.4) implies that (2.5) holds with $M \triangleq Q_2P_2$. Since $Q_2P_2 = P_2^{-1/2}(P_2^{1/2}Q_2P_2^{1/2})P_2^{1/2}$, M is positive semisimple. The rank conditions (2.12) follow from Sylvester's inequality. Expanding (A.2) and (A.3) yields

$$0 = AQ_1 + Q_1A^T + BVB^T, \tag{A.10}$$

$$0 = AQ_{12} + Q_{12}A_m^T + BVB_m^T, \tag{A.11}$$

$$0 = A_mQ_2 + Q_2A_m^T + B_mVB_m^T, \tag{A.12}$$

$$0 = A^TP_1 + P_1A + C^TRC, \tag{A.13}$$

$$0 = A^TP_{12} + P_{12}A_m - C^TRC_m, \tag{A.14}$$

$$0 = A_m^TP_2 + P_2A_m + C_m^TRC_m. \tag{A.15}$$

Since A_m is stable and (A_m,B_m) is controllable, standard results (e.g., [27], p. 277) imply that Q_2 is positive definite. Similarly, P_2 is positive definite.

It is easy to see at this point that A_m, B_m and C_m are independent of Q_1 and P_1 and thus (A.10) and (A.13) can be ignored. Now, substituting (2.10), (2.11) and the identities

$$Q_{12} = \hat{Q}\Gamma^T, \qquad P_{12} = -\hat{P}G^T, \tag{A.16}$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \qquad P_2 = G\hat{P}G^T, \tag{A.17}$$

into (A.11), (A.12), (A.14) and (A.15) yields

$$0 = A\hat{Q}\Gamma^T + \hat{Q}\Gamma^TA_m^T + BVB^T\Gamma^T, \tag{A.18}$$

$$0 = A_m\Gamma\hat{Q}\Gamma^T + \Gamma\hat{Q}\Gamma^TA_m^T + \Gamma BVB^T\Gamma^T, \tag{A.19}$$

$$0 = A^T \hat{P} G^T + \hat{P} G^T A_m + C^T R C G^T, \tag{A.20}$$

$$0 = A_m^T G \hat{P} G^T + G \hat{P} G^T A_m + G C^T R C G^T. \tag{A.21}$$

Computing (A.19)-Γ(A.18) implies

$$A_m = \Gamma A \hat{Q} \Gamma^T (\Gamma \hat{Q} \Gamma^T)^{-1}$$

which, since $\Gamma Q \Gamma^T = Q_2$, yields (2.9). Alternatively, (2.9) can be obtained from (A.21)-G(A.20).

If we now substitute (2.9) into (A.18)-(A.21) and use the easily verified relations (2.20), it follows that (A.19) = Γ(A.18) and (A.22) = G(A.21) and thus (A.19) and (A.21) are redundant. Finally, G^T(A.18)^T and (A.20)Γ yield (2.13) and (2.14), respectively. Note that these last multiplications entail no loss of generality since $\rho(G) = \rho(\Gamma) = n_m$.

To show that the optimal projection equations entail no loss of generality over (A.2)-(A.6), let $\hat{Q}, \hat{P}$ be extremal and define Q_{12}, Q_2, P_{12}, P_2 by (A.16) and (A.17) for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$ and let Q_1, P_1 satisfy (A.10) and (A.13). Then it is straightforward to reverse the steps taken in the proof to arrive at (A.2)-(A.6).□

<u>Proof of Proposition 2.1.</u> Extremal $\hat{Q}, \hat{P}$ leads to $\tilde{Q}, \tilde{P}$ as in (A.7) satisfying (A.2)-(A.6). Computing

$$\begin{aligned} J(A_m, B_m, C_m) &= \mathrm{tr}(Q_1 C^T R C - 2 Q_{12} C_m^T R C) + \mathrm{tr}(Q_2 C_m^T R C_m) \\ &= \mathrm{tr}[C^T R C (W_c - \hat{Q})], \end{aligned}$$

noting that (2.13), (2.14) are equivalent to (2.21), (2.22) because of (2.20) and using (2.23), leads to (2.15).□

References

1. D.A. Wilson, "Optimum solution of model-reduction problem", Proc. IEE, Vol. 117, pp. 1161-1165, 1970.

2. B.C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction", IEEE Trans. Autom. Contr., Vol. AC-26, pp. 17-32, 1981.

3. R.E. Skelton, "Cost Decomposition of linear systems with application to model reduction", Int. J. Contr., Vol. 32, pp. 1031-1055, 1980.

4. M. Aoki, "Control of large-scale dynamic systems by aggregation", IEEE Trans. Autom. Contr., Vol. AC-13, pp. 246-253, 1968.

5. D.A. Wilson, "Model reduction for multivariable systems", Int. J. Contr., Vol. 20, pp. 57-64, 1974.

6. R.N. Mishra and D.A. Wilson, "A new algorithm for optimal reduction of multivariable systems", Int. J. Contr., Vol. 31, pp. 443-466, 1980.

7. L. Pernebo and L.M. Silverman, "Model reduction via balanced state space representations", IEEE Trans. Autom. Contr., Vol. AC-27, pp. 382-387, 1982.

8. K.V. Fernando and H. Nicholson, "On the structure of balanced and other principal representations of SISO systems", IEEE Trans. Autom. Contr., Vol. AC-28, p. 228-231, 1983.

9. S. Shokoohi, L.M. Silverman and P.M. Van Dooren, "Linear time-variable systems: Balancing and model reduction", IEEE Trans. Autom. Contr., Vol. AC-28, pp. 810-822, 1983.

10. E.I. Verriest and T. Kailath, "On generalized balanced realizations", IEEE Trans. Autom. Contr., Vol. AC-28, pp. 833-844, 1983.

11. E.A. Jonckheere and L.M. Silverman, "A new set of invariants for linear systems - application to reduced-order compensator design", IEEE Trans. Autom. Contr., Vol AC-28, pp. 953-964, 1983.

12. R.E. Skelton and A. Yousuff, "Component cost analysis of large scale systems", Int. J. Contr., Vol. 37, pp. 285-304, 1983.

13. P.T. Kabamba, "Balanced gains and their significance for balanced model reduction", Proc. Conf. Inform. Sci. Sys., Princeton Univ., 1984.

14. C.R. Rao and S.K. Mitra, <u>Generalized Inverse of Matrices and its Applications</u>, John Wiley and Sons, New York, 1971.

15. S.L. Campbell and C.D. Meyer, Jr., <u>Generalized Inverses of Linear Transformations</u>, Pitman, London, 1979.

16. T. Kato, <u>Perturbation Theory for Linear Operators</u>, Springer-Verlag, New York, 1966.

17. C.T. Mullis and R.A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters", IEEE Trans. Circ. Syst., Vol. CAS-23, pp. 551-562, 1976.

18. D.C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters", AIAA 20th Aerospace Sciences Mtg., Orlando, Fl, Jan. 1982.

19. D.C. Hyland, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results", AIAA 21st Aerospace Sciences Mtg., Reno, NV, Jan. 1983.

20. D.C. Hyland and D.S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation", IEEE Conf. Dec. Contr., San Antonio, TX, Dec. 1983.

21. D.C. Hyland and D.S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation, IEEE Trans. Autom. Contr. (to appear).

22. D.C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection", Proc. AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

23. D.S. Bernstein and D.C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation", submitted for publication.

24. C.S. Sims, "Reduced-order modeling and filtering", in <u>Control and Dynamic Systems</u>, Vol. 18, pp. 55-103, 1982.

25. D.S. Bernstein and D.C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of infinite-dimensional systems", Proc. AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

26. M. Athans, "The matrix minimum principle", Inform. Contr., Vol. 11, pp. 592-606, 1968.

27. W.M. Wonham, <u>Linear Multivariable Control: A Geometric Approach</u>, Springer-Verlag, New York, 1974.

APPENDIX D

(REFERENCE [28])

THE OPTIMAL PROJECTION EQUATIONS FOR FINITE-DIMENSIONAL FIXED-ORDER DYNAMIC COMPENSATION OF INFINITE-DIMENSIONAL SYSTEMS*

by

Dennis S. Bernstein and David C. Hyland

ABSTRACT

One of the major difficulties in designing implementable finite-dimensional controllers for distributed parameter systems is that such systems are inherently infinite dimensional while controller dimension is severely constrained by on-line computing capability. While some approaches to this problem initially seek a correspondingly infinite-dimensional control law whose finite-dimensional approximation may be of impractically high order, the usual engineering approach involves first approximating the distributed parameter system with a high-order discretized model followed by design of a relatively low-order dynamic controller. Among the numerous approaches suggested for the latter step are model/controller reduction techniques used in conjunction with the standard LQG result. An alternative approach, developed in [36], relies upon the discovery that the necessary conditions for optimal fixed-order dynamic compensation can be transformed into a set of equations possessing remarkable structural coherence. The present paper generalizes this result to apply directly to the distributed parameter system itself. In contrast to the pair of operator Riccati equations for the "full-order" LQG case, the optimal finite-dimensional fixed-order dynamic compensator is characterized by four operator equations (two modified Riccati equations and two modified Lyapunov equations) coupled by an oblique projection whose rank is precisely equal to the order of the compensator and which determines the optimal compensator gains. This "optimal projection" is obtained by a full-rank factorization of the product of the finite-rank nonnegative-definite Hilbert-space operators which satisfy the pair of modified Lyapunov equations. The coupling represents a graphic portrayal of the demise of the classical separation principle for the finite-dimensional reduced-order controller case. The results obtained apply to a semigroup formulation in Hilbert space and thus are applicable to control problems involving a broad range of specific partial and functional differential equations.

---

*This work was performed at Lincoln Laboratory/MIT and was sponsored by the Department of the Air Force. Both authors are currently with Harris Corporation, Government Aerospace Systems Division, Controls Analysis and Synthesis Group, Melbourne, FL 32901.

1. INTRODUCTION

One of the major difficulties in designing active controllers for distributed parameter systems is that such systems are inherently infinite dimensional while implementable controllers are necessarily finite dimensional with controller dimension severely constrained by on-line computing capability. As pointed out by Balas ([1], see also [2]), control design for distributed parameter systems entails the practical constraints of 1) finitely many sensors and actuators, 2) a finite-dimensional controller and 3) natural system dissipation. The validity of 2) is apparent from the fact that processing and transmitting electrical signals by conventional analog or digital components constitutes finite-dimensional action. Although distributed parameter devices can also be utilized, their fabrication and implementation can incorporate at most a finite number of design specifications.* Hence, although distributed parameter systems are most accurately represented by infinite-dimensional models, real-world constraints require that implementable controllers be modelled as lumped parameter systems.

Clearly, the above observations effectively preclude the possibility of realizing infinite-dimensional controllers that involve full-state feedback or full-state estimation (see, e.g., [4-6] and the numerous references therein). Although finite-dimensional approximation schemes have been applied to optimal infinite-dimensional control laws ([7-9]), these results only guarantee optimality in the limit, i.e., as the order of the approximating controller increases without

*Examples of such components include tapped delay lines and surface acoustic wave devices. Although acoustoelectric convolvers ([3], p. 465) can perform continuous-time integration, synthesis of the desired impulse response kernel can incorporate only finitely many specified parameters. The obvious fact should also be noted that physical limitations impose an upper bound on the number of design parameters that can be incorporated in the construction of any device.

bound. Hence, there is no guarantee that a particular approximate (i.e., discretized) controller is actually optimal over the class of approximate controllers of a given order which may be dictated by implementation constraints. Moreover, even if an optimal approximate finite-dimensional controller could be obtained, it would almost certainly be suboptimal in the class of all controllers of the given order.

Although the usual engineering approach to this problem is to replace the distributed parameter system with a high-order finite-dimensional model, analogous, fundamental difficulties remain since application of LQG leads to a controller whose order is identical to that of the high-order approximate model. Attempts to remedy this problem usually rely upon some method of open-loop model reduction or closed-loop controller reduction (see, e.g., [10-15]). Most of these techniques (with the exception of [11]) are ad hoc in nature, however, and hence guarantees of optimality and stability may be lacking.

A more direct approach that avoids both model and controller reduction is to fix the controller structure and optimize the performance criterion with respect to the controller parameters. Although much effort has been devoted to this approach (see, e.g., [16-30]), progress in this direction has been impeded by the extreme complexity of the nonlinear matrix equations arising from the first-order necessary conditions. What was lacking, to quote the insightful remarks of [24], was a "deeper understanding of the structural coherence of these equations." The key to unlocking these unwieldy equations was subsequently discovered by Hyland in [31] and developed in [32-36]. Specifically, it was found that these equations harbored the definition of an oblique projection (i.e., idempotent matrix) which is a consequence of optimality and not the result of an ad hoc assumption. By exploiting the presence of this "optimal projection," the originally very complex stationary conditions can be transformed without loss of generality into much simpler and more tractable forms. The resulting equations

(see (2.10)-(2.17) of [36]) preserve the simple form of LQG relations for the gains in terms of covariance and cost matrices which, in turn, are determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations. This coupling, by means of the optimal projection, represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case. When, as a special case, the order of the compensator is required to be equal to the order of the plant, the modified Riccati equations immediately reduce to the standard LQG Riccati equations and the modified Lyapunov equations express the proviso that the compensator be minimal, i.e., controllable and observable. Since the LQG Riccati equations as such are nothing more than the necessary conditions for full-order compensation, the "optimal projection equations" appear to provide a clear and simple generalization of standard LQG theory.

The fact that the optimal projection equations consist of <u>four</u> coupled matrix equations, i.e., two modified Riccati equations and two modified Lyapunov equations, can readily be explained by the following simple reason. Reduced-order control-design methods often involve either LQG applied to a reduced-order model or model reduction applied to a full-order LQG design, and hence both approaches require the solution of precisely four equations: two Riccati equations (for LQG) plus two Lyapunov equations (for system reduction via balancing, as in [12,14]). The <u>coupled</u> form of the optimal projection equations is thus a strong reminder that the LQG and order-reduction operations <u>cannot</u> be iterated but must, in a certain sense, be performed simultaneously. This situation is partly due to the fact that the optimal projection matrix may not be of the form $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ even in the basis corresponding to the "balanced" realization ([12,14]). This point is explored in [37] where the solution to the optimal model-reduction problem is

characterized by a pair of modified Lyapunov equations which are also coupled by an optimal projection.

Returning now to the distributed parameter problem, it should be mentioned that notable exceptions to the previously-mentioned work on distributed parameter controllers are the contributions of Johnson ([38]) and Pearson ([39,40]) who suggest fixing the order of the finite-dimensional compensator while retaining the distributed parameter model. Progress in this direction, however, was impeded not only by the intractability of the optimality conditions that were available for the finite-dimensional problem (as in [16-30]), but also by the lack of a suitable generalization of these conditions to the infinite-dimensional case. The purpose of the present paper is to make significant progress in filling these gaps, i.e., by deriving explicit optimality conditions which directly characterize the optimal finite-dimensional fixed-order dynamic compensator for an infinite-dimensional system and which are exactly analogous to the highly-simplified optimal projection equations obtained in [31-34,36] for the finite-dimensional case. Specifically, instead of a system of four matrix equations we obtain a system of four <u>operator</u> equations whose solutions characterize the optimal finite-dimensional fixed-order dynamic compensator. Moreover, the optimal projection now becomes a bounded idempotent Hilbert-space operator whose rank is precisely equal to the order of the compensator.

The mathematical setting we use is standard: a linear time-invariant differential system in Hilbert space with additive white noise, finitely many controls and finitely many noisy measurements (thus satisfying the first practical constraint mentioned above). The input and output maps are assumed to be bounded. Since the only explicit assumption on the unbounded dynamics operator is that it generate a strongly continuous semigroup, the results are potentially applicable to a broad range of specific partial and functional differential equations. The <u>actual</u> applicability of our results is essentially limited by

practical constraint 3). Since we are concerned with the steady-state problem, we implicitly assume that the distributed parameter system is stabilizable, i.e., that there exists a dynamic compensator of a given order such that the closed-loop system is uniformly stable. We note that stabilizing compensators do exist for the wide class of problems considered in [41] and [42] which includes delay, parabolic and damped hyperbolic systems. The question of how much damping is required for stabilizability of hyperbolic systems is a crucial issue in designing controllers for large flexible space structures ([7, 43-49a]).

It is important to point out that the results of this paper can immediately be specialized to finite-dimensional systems by requiring that the Hilbert space characterizing the dynamical system be finite-dimensional. Then all unboundedness considerations can be ignored, adjoints can be interpreted as transposes and other obvious simplifications can be invoked. The only mathematical aspect requiring attention is the treatment of white noise which, for general handling of the infinite-dimensional case, is interpreted according to [6].* For the finite-dimensional case, however, the standard classical notions suffice and the results go through with virtually no modifications.

The contents of the paper are as follows. Section 2 contains preliminary notation in addition to particular results for use later in the paper. Section 3 presents the optimal steady-state finite-dimensional fixed-order dynamic-compensation problem and the Main Theorem gives the necessary conditions

---

*Alternatively, we could have adopted the white noise formulation of [4]. However, this would have required additional technical assumptions on the plant (see, Theorem 5.35, p. 148 of [4]). The main difference between the two white noise formalisms is that Balakrishnan works with finitely additive rather than countably additive measures. Strictly speaking, then, even in finite dimensions Balakrishnan's white noise is different from the standard notion (see [6], pp. 307, 315).

in the form of the optimal projection equations (3.15)-(3.18). We then develop a series of results which serve to elucidate several aspects of the Main Theorem. Section 4 is devoted to the proof of the Main Theorem. The reader is alerted to the two crucial steps required. The first step involves generalizing to the infinite-dimensional case the derivation of the necessary conditions in their "primitive" form (see (4.27)-(4.29) and (4.48)-(4.53)). The derivation in [31-33, 36] involving Lagrange multipliers is invalid in the infinite-dimensional case due to the presence of the unbounded system-dynamics operator. Instead, we use the gramian form of the closed-loop covariance operator to obtain a dual problem formulation and then proceed to derive the primitive necessary conditions by means of a lengthy, but direct, computation (Lemma 4.7). The second crucial step involves transformation of the primitive form of the necessary conditions to the final form given in the Main Theorem. This laborious computation was first carried out in [31,32] and was subsequently facilitated in [33,36] by means of a judicious change of variables (see (4.32), (4.33)). Finally, some concluding remarks are given in Section 5.

2. PRELIMINARIES

In this section we introduce general notation along with basic definitions and results for use in later sections. Our principal references are [6], [50] and [51].

Throughout this section let $\underline{H}$, $\underline{H}'$ and $\underline{H}''$ denote real separable Hilbert spaces with norm $||\cdot||$ and inner product $<\cdot,\cdot>$ and let $\underline{B}(\underline{H},\underline{H}')$ denote the space of bounded linear operators from $\underline{H}$ into $\underline{H}'$. For $L \in \underline{B}(\underline{H},\underline{H}')$, $||L||$ is the norm of L, $\underline{R}(L)$ is the range of L, $\underline{N}(L)$ is the null space of L, $\rho(L)$ is the rank of L (set $\rho(L) = \infty$ if L does not have finite rank), L^{-1} is the inverse of L when L is invertible, i.e., when L has a bounded inverse, L^* is the adjoint of L and $L^{-*} \triangleq (L^*)^{-1}$. Recall that $||L|| = ||L^*||$ and that $\rho(L) = \rho(L^*)$ ([50], p. 161). If $LL^* = L^*L$ then L is normal. Now suppose that $\underline{H} = \underline{H}'$ so that $L \in \underline{B}(\underline{H}) \triangleq \underline{B}(\underline{H},\underline{H})$. If $L = L^*$ then L is selfadjoint. If L is selfadjoint and $<Lx,x> \geq 0$, $x \in \underline{H}$, then L is nonnegative definite. Note that the selfadjointness assumption is included in the definition since the Hilbert spaces are assumed real. If L is nonnegative definite then $L^{\frac{1}{2}}$ denotes the (unique) nonnegative-definite square root of L. Call L semisimple (resp., real semisimple, nonnegative semisimple) if there exists invertible $S \in \underline{B}(\underline{H})$ such that SLS^{-1} is normal (resp., selfadjoint, nonnegative definite). This implies that SLS^{-1} has a complete set of orthonormal eigenvectors and, in the real-semisimple or nonnegative-semisimple cases, has real or nonnegative eigenvalues.

Recall that if $S \in \underline{B}(\underline{H})$ is compact then S has at most a countable number of eigenvalues and all nonzero eigenvalues have finite multiplicity. Hence, for $L \in \underline{B}(\underline{H},\underline{H}')$ compact, let $\{\alpha_i\}$ be the (at most countable) sequence of eigenvalues of $(LL^*)^{\frac{1}{2}}$ with appropriate multiplicity and $\alpha_1 \geq \alpha_2 \geq \ldots > 0$ ([50], p. 261). Then $\underline{B}_1(\underline{H},\underline{H}')$ denotes the set of trace class (or nuclear) operators, i.e., the set of L for which $\sum_i \alpha_i < \infty$ ([50], p. 521). $\underline{B}_1(\underline{H},\underline{H}')$ is a Banach space

with norm

$$||L||_1 \triangleq \sum_i \alpha_i .$$

If $\sum_i \alpha_i^2 < \infty$ then $L \in \underline{B}_2(\underline{H},\underline{H}')$, the set of Hilbert-Schmidt operators, which is a Banach space with norm

$$||L||_2 \triangleq [\sum_i \alpha_i^2]^{\frac{1}{2}} .$$

Note that $||L|| \leq ||L||_2 \leq ||L||_1$, $||L|| = ||L^*||$, $||L||_1 = ||L^*||_1$ and $||L||_2 = ||L^*||_2$. If $\underline{H} = \underline{H}'$, then we write $\underline{B}_1(\underline{H})$ and $\underline{B}_2(\underline{H})$ for $\underline{B}_1(\underline{H},\underline{H})$, and $\underline{B}_2(\underline{H},\underline{H})$, respectively. Note that if nonnegative-definite $L \in \underline{B}_1(\underline{H})$ then $L^{\frac{1}{2}} \in \underline{B}_2(\underline{H})$.

If $L \in \underline{B}_1(\underline{H},\underline{H}')$ and $S \in \underline{B}(\underline{H}',\underline{H}'')$ then

$$||SL||_1 \leq ||S|| \; ||L||_1$$

and hence $SL \in \underline{B}_1(\underline{H},\underline{H}')$. Similarly, under suitable hypotheses,

$$||LS||_1 \leq ||S|| \; ||L||_1,$$

and

$$||SL||_1 \leq ||S||_2 \; ||L||_2.$$

Lemma 2.1. Suppose $L \in \underline{B}_1(\underline{H})$ and let $\{\lambda_i\}$ denote the nonzero eigenvalues of L with appropriate multiplicity. Then ([51], p. 89)

$$\sum_i |\lambda_i| \leq ||L||_1 .$$

If L is selfadjoint then ([50], p. 522)

$$\sum_i |\lambda_i| = ||L||_1 .$$

If L is nonnegative definite then

$$\sum_i \lambda_i = ||L||_1 .$$

Let $L \in \underline{B}_1(\underline{H})$. Then define ([50], p. 523) the trace functional $\text{tr}: \underline{B}_1(\underline{H}) \to \mathbb{R}$ by

$$\text{tr } L \triangleq \sum_i \langle L\,\phi_i, \phi_i \rangle,$$

where the summation is independent of the choice of orthonormal basis $\{\phi_i\}$. The trace satisfies $\text{tr } L = \text{tr } L^*$, $\text{tr } SL = \text{tr } LS$ for all $S \in \underline{B}(\underline{H})$, $\text{tr } ST = \text{tr } TS$ for all $S,T \in \underline{B}_2(\underline{H})$ and $\text{tr}(\alpha T + \beta S) = \alpha(\text{tr } T) + \beta(\text{tr } S)$ for all $\alpha,\beta \in \mathbb{R}$ and $S,T \in \underline{B}_1(\underline{H})$.

Lemma 2.2. Suppose $L \in \underline{B}_1(\underline{H})$ and let $\{\lambda_i\}$ denote the nonzero eigenvalues of L with appropriate multiplicity. Then ([51], p. 139)

$$\text{tr } L = \sum_i \lambda_i$$

and hence

$$|\text{tr } L| \le ||L||_1.$$

If L is nonnegative definite then

$$\text{tr } L = ||L||_1.$$

Corollary 2.1. For each $S \in \underline{B}(\underline{H})$ the linear functionals

$$L \to \text{tr } SL: \underline{B}_1(\underline{H}) \to \mathbb{R},$$

$$L \to \text{tr } LS: \underline{B}_1(\underline{H}) \to \mathbb{R}$$

are continuous. For each $L \in \underline{B}_1(\underline{H})$ the linear functionals

$$S \to \text{tr } LS: \underline{B}(\underline{H}) \to \mathbb{R},$$

$$S \to \text{tr } SL: \underline{B}(\underline{H}) \to \mathbb{R}$$

are continuous.

Although showing that a bounded linear operator is trace class is slightly more involved than the above characterizations of $\underline{B}_1(\underline{H})$, the following result will suffice for our purposes (see [52], p. 96, or [6], p. 115).

Lemma 2.3. Let $L \in \underline{B}(\underline{H})$ be nonnegative definite. Then

$$\sum_i \langle L\,\phi_i,\ \phi_i \rangle ,$$

whether finite or infinite, is independent of the orthonormal basis $\left\{\phi_i\right\}$. The summation is finite if and only if $L \in \underline{B}_1(\underline{H})$.

Many of the operators introduced in the following section have finite-dimensional domain or range space and hence are degenerate, i.e., have finite rank. Recall that degenerate operators are necessarily trace class. The following result, which generalizes Theorem 2.1, p. 240 of [53] in certain respects, will be fundamental in decomposing finite-rank operators.

Lemma 2.4. Suppose $L_1,\ldots,L_r \in B(\underline{H},\underline{H}')$ have finite rank. Then there exists a finite-dimensional subspace $\underline{M} \subset \underline{H}$ such that $L_i\underline{M}^\perp = 0, i=1,\ldots,r$. Furthermore, if $\underline{H} = \underline{H}'$ then $\underline{M}$ can be chosen such that $L_i\underline{M} \subset \underline{M}, i=1,\ldots,r$.

Proof. It suffices to consider the case r=1. Writing L for L_1, note that since $\rho(L^*)<\infty$, $\underline{N}(L)^\perp = \underline{R}(L^*)$ ([50], p. 155) and $\underline{N}(L)$ is closed, the first statement holds with $\underline{M} = \underline{N}(L)^\perp$. When $\underline{H} = \underline{H}'$ set $\underline{M} = \underline{N}(L)^\perp + \underline{R}(L)$ and note that $\underline{M}^\perp = \underline{N}(L) \cap \underline{R}(L)^\perp \subset \underline{N}(L)$ and $L\underline{M} \subset \underline{R}(L) \subset \underline{M}$. □

The following generalization of Sylvester's inequality ([54], p. 66) will be used repeatedly in handling finite-rank operators.

Lemma 2.5. Let $L \in \underline{B}(\underline{H},\underline{H}')$ and $S \in \underline{B}(\underline{H}',\underline{H}'')$. Then

$$\rho(SL) \leq \min\left\{\rho(S),\ \rho(L)\right\}. \tag{2.1}$$

If $\dim \underline{H}' = \nu < \infty$, then

$$\rho(S) + \rho(L) - \nu \leq \rho(SL). \tag{2.2}$$

Proof. If either S or L does not have finite rank then (2.1) is immediate. If both S and L have finite rank then the standard arguments ([54]) used to prove the finite-dimensional version of (2.1) remain valid. To prove

(2.2), note that Lemma 2.4 implies that there exist orthonormal bases for $\underline{H}$ and $\underline{H}'$ with respect to which L has the matrix representation $[\tilde{L} \;\; 0]$, where $\tilde{L} \in \mathbb{R}^{\nu \times p}$. Similarly, there exist orthonormal bases for $\underline{H}'$ and $\underline{H}''$ with respect to which S has the matrix representation $\begin{bmatrix} \tilde{S} \\ 0 \end{bmatrix}$, where $\tilde{S} \in \mathbb{R}^{q \times \nu}$. Since the two cited bases for $\underline{H}'$ may be different, let orthogonal $U \in \mathbb{R}^{\nu \times \nu}$ be the matrix representation (with respect to either basis for $\underline{H}'$) for the change in orthonormal basis ([6], p. 100). Hence SL has the matrix representation $\begin{bmatrix} \tilde{S}U\tilde{L} & 0 \\ 0 & 0 \end{bmatrix}$ and (2.2) follows from the known result ([54], p. 66). □

As in the proof of Lemma 2.5, we shall utilize the infinite-matrix representation of an operator with respect to an orthonormal basis. All matrix representations given here will consist of real entries since the Hilbert spaces involved are real. When the orthonormal bases are specified and no confusion can arise, we shall not differentiate between an operator and its matrix representation. We shall use the infinite identity matrix I_∞ interchangeably with the identity $I_{\underline{H}}$ on $\underline{H}$.

When dealing with finite-dimensional Euclidean spaces the notation and terminology introduced above will be utilized with only minor changes. For example, bounded linear operators will be represented by matrices whose elements are determined according to fixed orthonormal bases and hence we identify $\mathbb{R}^{m \times n} = \underline{B}(\mathbb{R}^n, \mathbb{R}^m)$. Note that if $L \in \underline{B}(\mathbb{R}^n, \underline{H})$ and $S \in \underline{B}(\underline{H}, \mathbb{R}^m)$ then SL is an mxn matrix which is independent of any particular orthonormal basis for $\underline{H}$. The transposes of $x \in \mathbb{R}^n \triangleq \mathbb{R}^{n \times 1}$ and $M \in \mathbb{R}^{m \times n}$ are denoted by x^T and M^T and $M^{-T} \triangleq (M^T)^{-1}$. Let I_n denote the nxn identity matrix.

To specialize some of the above operator terminology to matrices, let $M \in \mathbb{R}^{n \times n}$. We shall say M is <u>nonnegative</u> (resp., <u>positive</u>) <u>diagonal</u> if M is diagonal with nonnegative (resp. positive) diagonal elements. M is <u>nonnegative</u> (resp., <u>positive</u>) <u>definite</u> if M is symmetric and $x^T M x \geq 0$ (resp., $x^T M x > 0$), $x \in \mathbb{R}^n$. Recall that M is symmetric (resp., nonnegative definite, positive

definite) if and only if there exists orthogonal $U \in \mathbb{R}^{n\times n}$ such that UMU^T is diagonal (resp., nonnegative diagonal, positive diagonal). M is <u>semisimple</u> ([55], p. 13), or <u>nondefective</u> ([56], p. 375), if M has n linearly independent eigenvectors, i.e., M has a diagonal Jordan canonical form over the complex field. M is <u>real</u> (resp., <u>nonnegative</u>, <u>positive</u>) <u>semisimple</u> if M is semisimple with real (resp., nonnegative, positive) eigenvalues. Note that M is real (resp., nonnegative, positive) semisimple if and only if there exists invertible $S \in \mathbb{R}^{n\times n}$ such that SMS^{-1} is diagonal (resp., nonnegative diagonal, positive diagonal). Alternatively, M is real (resp., nonnegative, positive) semisimple if and only if there exists invertible $S \in \mathbb{R}^{n\times n}$ such that SMS^{-1} is symmetric (resp., nonnegative definite, positive definite).

<u>Lemma 2.6</u>. The product of two nonnegative- (resp., positive-) definite matrices is nonnegative (resp., positive) semisimple.

<u>Proof</u>. If $S, L \in \mathbb{R}^{n\times n}$ are both nonnegative (resp., positive) definite then by Theorem 6.2.5, p. 123 of [55] there exists invertible $\phi \in \mathbb{R}^{n\times n}$ such that $D_S \triangleq \phi^{-1}S\phi^{-T}$ and $D_L \triangleq \phi^T L\phi$ are nonnegative (resp., positive) diagonal. Hence, $SL = \phi D_S D_L \phi^{-1}$ is nonnegative (resp., positive) semisimple, as desired. Alternatively, if either S or L is positive definite, then the result follows from $SL = L^{-1/2}(L^{1/2}SL^{1/2})L^{1/2}$ if L is positive definite or $SL = S^{1/2}(S^{1/2}LS^{1/2})S^{-1/2}$ if S is positive definite. $\square$

3. PROBLEM STATEMENT AND THE MAIN THEOREM

We consider the following steady-state fixed-order dynamic-compensation problem. Given the dynamical system on $[0, \infty)$

$$\dot{x}(t) = Ax(t) + Bu(t) + H_1 w(t), \tag{3.1}$$

$$y(t) = Cx(t) + H_2 w(t), \tag{3.2}$$

design a finite-dimensional fixed-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \tag{3.3}$$

$$u(t) = C_c x_c(t) \tag{3.4}$$

which minimizes the steady-state performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \mathbb{E}[\langle R_1 x(t), x(t)\rangle + u(t)^T R_2 u(t)]. \tag{3.5}$$

The following data are assumed. The state $x(t)$ is an element of a real separable Hilbert space $\underline{H}$ and the state differential equation is interpreted in the weak sense (see, e.g., [6], pp. 229, 317). The closed, densely defined operator A: $\underline{D}(A) \subset \underline{H} \to \underline{H}$ generates a strongly continuous semigroup e^{At}, $t \geq 0$. The control $u(t) \in \mathbb{R}^m$, $B \in \underline{B}(\mathbb{R}^m, \underline{H})$ and the operator $R_1 \in \underline{B}_1(\underline{H})$ and the matrix $R_2 \in \mathbb{R}^{m \times m}$ are nonnegative definite and positive definite, respectively. $w(\cdot)$ is a zero-mean Gaussian "standard white noise process" in $L_2((0,\infty), \underline{H}')$ (see [6], p. 314), where H' is a real separable Hilbert space, $H_1 \in \underline{B}_2(\underline{H}', \underline{H})$, $H_2 \in \underline{B}(\underline{H}', \mathbb{R}^\ell)$ and "$\mathbb{E}$" denotes expectation. We assume that $H_1 H_2^* = 0$, i.e., the disturbance and measurement noises are independent, and that $V_2 \triangleq H_2 H_2^* \in \mathbb{R}^\ell$ is positive definite, i.e., all measurements are noisy. Note that $V_1 \triangleq H_1 H_1^* \in \underline{B}_1(\underline{H})$ is nonnegative

definite and trace class.* The initial state $x(0)$ is Gaussian and independent of $w(\cdot)$. The observation $y(t) \in \mathbb{R}^{\ell}$ and $C \in \underline{B}(\underline{H},\mathbb{R}^{\ell})$. The dimension of the compensator state $x_c(t)$ is of fixed, finite order $n_c \leq \dim \underline{H}$ and the optimization is performed over $A_c \in \mathbb{R}^{n_c x n_c}$, $B_c \in \mathbb{R}^{n_c x \ell}$ and $C_c \in \mathbb{R}^{m x n_c}$.

To handle the closed-loop system (3.1)-(3.4), we introduce the augmented state space $\tilde{\underline{H}} \triangleq \underline{H} \oplus \mathbb{R}^{n_c}$ which is a real separable Hilbert space with inner product $\langle \tilde{x}_1, \tilde{x}_2 \rangle \triangleq \langle x_1, x_2 \rangle + x_{c1}^T x_{c2}$, $\tilde{x}_i \triangleq (x_i, x_{ci})$. An operator $L \in \underline{B}(\tilde{\underline{H}})$ has a "decomposition" into operators $L_1 \in \underline{B}(\underline{H})$, $L_{12} \in \underline{B}(\mathbb{R}^{n_c},\underline{H})$, $L_{21} \in \underline{B}(\underline{H},\mathbb{R}^{n_c})$ and $L_2 \in \mathbb{R}^{n_c x n_c}$ in the sense that for $\tilde{x} \triangleq (x,x_c) \in \tilde{\underline{H}}$, $L\tilde{x} = (L_1 x + L_{12} x_c, L_{21} x + L_2 x_c)$, or, in "block" form,

$$L = \begin{bmatrix} L_1 & L_{12} \\ L_{21} & L_2 \end{bmatrix} .$$

For later use note that

$$||L|| \leq ||L_1|| + ||L_{12}|| + ||L_{21}|| + ||L_{22}||$$

and

$$L^* = \begin{bmatrix} L_1^* & L_{21}^* \\ L_{12}^* & L_2^T \end{bmatrix} .$$

We can similarly construct unbounded operators in $\tilde{\underline{H}}$. Hence, define the closed-loop dynamics operator $\tilde{A}$: $\underline{D}(\tilde{A}) \subset \tilde{\underline{H}} \rightarrow \tilde{\underline{H}}$ on the dense domain $\underline{D}(\tilde{A}) \triangleq \underline{D}(A) \times \mathbb{R}^{n_c}$ by $\tilde{A}\tilde{x} = (Ax + BC_c x_c, B_c Cx + A_c x_c)$. Since $\tilde{A}$ can be represented by

$$\tilde{A} = \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & BC_c \\ B_c C & A_c \end{bmatrix}$$

---

*We must require that R_1 and V_1 be nuclear since covariance operators in the white noise formulation of [6] are not necessarily nuclear as they are in the formulation of [4].

and since the closed-loop operator $\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$: $\underline{D}(\widetilde{A}) \rightarrow \widetilde{\underline{H}}$ generates the strongly continuous semigroup $\begin{bmatrix} e^{At} & 0 \\ 0 & I_{n_c} \end{bmatrix}$, $t \geq 0$, it follows from Theorem 2.1, p. 497 of [50] that $\widetilde{A}$ is also closed and generates a strongly continuous semigroup $e^{\widetilde{A}t} \in \underline{B}(\widetilde{\underline{H}})$, $t \geq 0$. To guarantee that J is finite and independent of initial conditions we restrict our attention to the set of admissible stabilizing compensators

$$\underline{A} \triangleq \left\{(A_c, B_c, C_c): \; e^{\widetilde{A}t} \text{ is exponentially stable}\right\}.$$

Hence if $(A_c, B_c, C_c) \in \underline{A}$ then there exist $M > 0$ and $\beta > 0$ such that

$$||e^{\widetilde{A}t}|| \leq Me^{-\beta t}, \; t \geq 0. \tag{3.6}$$

Since the value of J is independent of the internal realization of the compensator, we can further restrict our attention to

$$\underline{A}_+ \triangleq \left\{(A_c, B_c, C_c) \in \underline{A}: \; (A_c, B_c) \text{ is controllable and } (C_c, A_c) \text{ is observable}\right\}.$$

The following lemma is required for the statement of the Main Theorem.

Lemma 3.1. Suppose $\hat{Q}$, $\hat{P} \in \underline{B}(\underline{H})$ have finite rank and are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_c$ then there exist $G, \Gamma \in \underline{B}(\underline{H}, \mathbb{R}^{n_c})$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\hat{Q}\hat{P} = G^* M \Gamma, \tag{3.7}$$

$$\Gamma G^* = I_{n_c}. \tag{3.8}$$

Proof. By Lemma 2.4 there exists a finite-dimensional subspace $\underline{M} \subset \underline{H}$ such that $\hat{Q}\underline{M} \subset \underline{M}$, $\hat{Q}\underline{M}^{\perp} = 0$, $\hat{P}\underline{M} \subset \underline{M}$ and $\hat{P}\underline{M}^{\perp} = 0$. Hence there exists an orthonormal basis for $\underline{H}$ with respect to which $\hat{Q}$ and $\hat{P}$ have the infinite-matrix representations

$\hat{Q} = \begin{bmatrix} \hat{Q}_1 & 0 \\ 0 & 0 \end{bmatrix}$, $\hat{P} = \begin{bmatrix} \hat{P}_1 & 0 \\ 0 & 0 \end{bmatrix}$, where $\hat{Q}_1, \hat{P}_1 \in \mathbb{R}^{r\times r}$ are nonnegative definite and $r \triangleq \dim \underline{M}$. Since by Lemma 2.6 there exists invertible $\Psi \in \mathbb{R}^{r\times r}$ such that $\tilde{\Lambda} = \Psi^{-1}\hat{Q}_1\hat{P}_1\Psi$ is nonnegative diagonal, we have

$$\hat{Q}\hat{P} = \begin{bmatrix} \Psi & 0 \\ 0 & I_\infty \end{bmatrix} \begin{bmatrix} \tilde{\Lambda} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Psi^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

which shows that $\hat{Q}\hat{P}$ is nonnegative semisimple. If, furthermore, $\rho(\hat{Q}\hat{P}) = n_c$ then it is clear that Ψ can be chosen (i.e., modified by an orthogonal matrix) so that $\tilde{\Lambda} = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}$, where $\Lambda \in \mathbb{R}^{n_c \times n_c}$ is positive diagonal. Hence,

$$\hat{Q}\hat{P} = \begin{bmatrix} \Psi & 0 \\ 0 & I_\infty \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_{n_c} \\ 0 \end{bmatrix} \\ 0 \end{bmatrix} \Lambda \begin{bmatrix} \begin{bmatrix} I_{n_c} & 0 \end{bmatrix} & 0 \end{bmatrix} \begin{bmatrix} \Psi^{-1} & 0 \\ 0 & I_\infty \end{bmatrix},$$

which shows that (3.7) and (3.8) are satisfied with

$$G = \begin{bmatrix} \begin{bmatrix} S^T & 0 \end{bmatrix} & 0 \end{bmatrix} \begin{bmatrix} \Psi^T & 0 \\ 0 & I \end{bmatrix}, \quad M = S^{-1}\Lambda S, \quad \Gamma = \begin{bmatrix} \begin{bmatrix} S^{-1} & 0 \end{bmatrix} & 0 \end{bmatrix} \begin{bmatrix} \Psi^T & 0 \\ 0 & I_\infty \end{bmatrix},$$

for all invertible $S \in \mathbb{R}^{n_c \times n_c}$. □

We shall refer to $G, \Gamma \in \underline{B}(\underline{H}, \mathbb{R}^{n_c})$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ satisfying (3.7) and (3.8) as a (G,M,Γ)-factorization of $\hat{Q}\hat{P}$. For convenience in stating the Main Theorem define

$$\Sigma \triangleq BR_2^{-1}B^*, \quad \bar{\Sigma} \triangleq C^*V_2^{-1}C.$$

Main Theorem. Suppose $(A_c, B_c, C_c) \in \underline{A}_+$ solves the steady-state fixed-order dynamic-compensation problem. Then ther exist nonnegative-definite Q, P, $\hat{Q}$, $\hat{P} \in \underline{B}_1(\underline{H})$ such that A_c, B_c and C_c are given by

$$A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P)G^*, \tag{3.9}$$

$$B_c = \Gamma Q C^* V_2^{-1}, \tag{3.10}$$

$$C_c = -R_2^{-1}B^*PG^*, \tag{3.11}$$

for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$, and such that with $\tau \triangleq G^*\Gamma$ the following conditions are satisfied:

$$Q: \ \underline{D}(A^*) \to \underline{D}(A), \quad P: \ \underline{D}(A) \to \underline{D}(A^*), \tag{3.12a,b}$$

$$\hat{Q}: \ \underline{H} \to \underline{D}(A), \quad \hat{P}: \ \underline{H} \to \underline{D}(A^*), \tag{3.13a,b}$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_c, \tag{3.14a,b,c)*}$$

$$0 = (A - \tau Q\bar{\Sigma})Q + Q(A - \tau Q\bar{\Sigma})^* + V_1 + \tau Q\bar{\Sigma}Q\tau^*, \tag{3.15}$$

$$0 = (A - \Sigma P\tau)^*P + P(A - \Sigma P\tau) + R_1 + \tau^*P\Sigma P\tau, \tag{3.16}$$

$$0 = [(A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^* + Q\bar{\Sigma}Q]\tau^*, \tag{3.17}$$

$$0 = [(A - Q\bar{\Sigma})^*\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P]\tau. \tag{3.18}$$

---

*(3.14a) refers to $\rho(\hat{Q}) = n_c$, etc.

The content of the Main Theorem is clearly a set of necessary conditions which characterize the optimal steady-state fixed-order dynamic compensator when it exists. These necessary conditions consist of a system of four operator equations including a pair of modified Riccati equations (3.15) and (3.16) and a pair of modified Lyapunov equations (3.17) and (3.18). The salient feature of these four equations is the coupling by the operato $\tau \in \underline{B}(\underline{H})$ which, because of (3.8), is idempotent, i.e., $\tau^2 = \tau$. In general, τ is an <u>oblique</u> projection and not an orthogonal projection since there is no requirement that τ be selfadjoint. Additional features of the Main Theorem will be discussed in the remainder of this section. For convenience, let G, M, Γ, τ, Q, P, $\hat{Q}$ and $\hat{P}$ be as given by the Main Theorem and define $\Lambda \triangleq \mathrm{diag}\,(\lambda_1, \ldots, \lambda_{n_c})$, where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{n_c} > 0$ are the eigenvalues of M.

We begin by noting that if x_c is replaced by Sx_c, where $S \in \mathbb{R}^{n_c x n_c}$ is invertible, then an "equivalent" compensator is obtained with (A_c, B_c, C_c) replaced by $(SA_cS^{-1}, SB_c, C_cS^{-1})$.

<u>Proposition 3.1</u>. Let $(A_c, B_c, C_c) \in \underline{A}_+$. If $S \in \mathbb{R}^{n_c x n_c}$ is invertible then $(SA_cS^{-1}, SB_c, C_cS^{-1}) \in \underline{A}_+$ and

$$J(A_c, B_c, C_c) = J(SA_cS^{-1}, SB_c, C_cS^{-1}). \tag{3.19}$$

<u>Proof</u>. Although the result is obvious from system-theoretic arguments, we shall prove it analytically by utilizing elements of the development in Section 4. Define $\tilde{S} \triangleq \begin{bmatrix} I_\infty & 0 \\ 0 & S \end{bmatrix} \in \underline{B}(\tilde{\underline{H}})$ and note that replacing (A_c, B_c, C_c) by $(SA_cS^{-1}, SB_c, C_cS^{-1})$ is equivalent to replacing $\tilde{A}$, $\tilde{V}$ and $\tilde{R}$ by $\tilde{S}\tilde{A}\tilde{S}^{-1}$, $\tilde{S}\tilde{V}\tilde{S}^*$ and $\tilde{S}^{-*}\tilde{R}\tilde{S}^{-1}$, respectively. If $M, \beta > 0$ satisfy (3.6) then a straightforward application of the Hille-Yosida Theorem ([57], pp. 153-5) shows that the strongly continuous semigroup generated by $\tilde{S}\tilde{A}\tilde{S}^{-1}$ satisfies $||e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t}|| \leq ||\tilde{S}||\ ||\tilde{S}^{-1}||Me^{-\beta t}$, which

proves the first assertion. Since $\tilde{S}e^{\tilde{A}t}\tilde{S}^{-1}$, $t \geq 0$, is also a strongly continuous semigroup with generator $\tilde{S}\tilde{A}\tilde{S}^{-1}$, it follows that $\tilde{S}e^{\tilde{A}t}\tilde{S}^{-1} = e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t}$. Hence

$$\int_0^{\infty} e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t}(\tilde{S}\tilde{V}\tilde{S}^{*})e^{(\tilde{S}\tilde{A}\tilde{S}^{-1})^{*}t}dt = \tilde{S}\tilde{Q}\tilde{S}^{*}$$

and (3.19) follows from $\mathrm{tr}\,\tilde{Q}\tilde{R} = \mathrm{tr}\,(\tilde{S}\tilde{Q}\tilde{S}^{*})(\tilde{S}^{-*}\tilde{R}\tilde{S}^{-1})$. $\square$

In view of Proposition 3.1 one would expect the Main Theorem to apply also to $(SA_cS^{-1}, SB_c, C_cS^{-1})$. Indeed, it may be noted that no claim was made as to the uniqueness of the (G,M,Γ)-factorization of $\hat{\hat{Q}}\hat{\hat{P}}$ used to determine A_c, B_c and C_c in (3.9)-(3.11). These observations are reconciled by the following result which shows that a transformation of the compensator state basis corresponds to the alternative factorization $\hat{\hat{Q}}\hat{\hat{P}} = (S^{-T}G)^{T}(SMS^{-1})(S\Gamma)$ and, moreover, that all (G,M,Γ)-factorizations of $\hat{\hat{Q}}\hat{\hat{P}}$ are related by a nonsingular transformation. Note that τ remains invariant over the class of factorizations.

<u>Proposition 3.2</u>. If $S \in \mathbb{R}^{n_c \times n_c}$ is invertible then $\overline{G} \triangleq S^{-T}G$, $\overline{\Gamma} \triangleq S\Gamma$ and $\overline{M} \triangleq SMS^{-1}$ satisfy

$$\hat{\hat{Q}}\hat{\hat{P}} = \overline{G}^{*}\overline{M}\overline{\Gamma}, \tag{3.7)'}$$

$$\overline{\Gamma}\overline{G}^{*} = I_{n_c}. \tag{3.8)'}$$

Conversely, if $\overline{G}, \overline{\Gamma} \in \underline{B}(\underline{H}, \mathbb{R}^{n_c})$ and invertible $\overline{M} \in \mathbb{R}^{n_c \times n_c}$ satisfy (3.7)' and (3.8)', then there exists invertible $S \in \mathbb{R}^{n_c \times n_c}$ such that $\overline{G} = S^{-T}G$, $\overline{\Gamma} = S\Gamma$ and $\overline{M} = SMS^{-1}$.

Proof. The first part of the proposition is immediate. The second part follows by taking $S \triangleq \overline{M}^{-1}\overline{\Gamma}G^{*}M^{-1}$, noting $S^{-1} = M\overline{\Gamma}\overline{G}^{*}\overline{M}^{-1}$ and using the identities $\overline{\Gamma}G^{*}M\Gamma\overline{G}^{*} = M$ and $M\Gamma\overline{G}^{*} = \Gamma\overline{G}^{*}\overline{M}$. □

The next result shows that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ.

Proposition 3.3. There exists invertible $\Phi \in \underline{B}(\underline{H})$ such that

$$\hat{Q} = \Phi^{-1}\begin{bmatrix}\Lambda_{\hat{Q}} & 0\\ 0 & 0\end{bmatrix}\Phi^{-*}, \qquad \hat{P} = \Phi^{*}\begin{bmatrix}\Lambda_{\hat{P}} & 0\\ 0 & 0\end{bmatrix}\Phi, \tag{3.20a,b}$$

$$\hat{Q}\hat{P} = \Phi^{-1}\begin{bmatrix}\Lambda & 0\\ 0 & 0\end{bmatrix}\Phi, \qquad \tau = \Phi^{-1}\begin{bmatrix}I_{n_c} & 0\\ 0 & 0\end{bmatrix}\Phi, \tag{3.21a,b}$$

where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \mathbb{R}^{n_c \times n_c}$ are positive diagonal and $\Lambda_{\hat{Q}}\Lambda_{\hat{P}} = \Lambda$. Consequently,

$$\hat{Q} = \tau\hat{Q}, \qquad \hat{P} = \hat{P}\tau. \tag{3.22a,b}$$

Proof. Proceeding as in the proof of Lemma 3.1, choose an orthonormal basis for $\underline{H}$ with respect to which $\hat{Q} = \begin{bmatrix}\hat{Q}_1 & 0\\ 0 & 0\end{bmatrix}$ and $\hat{P} = \begin{bmatrix}\hat{P}_1 & 0\\ 0 & 0\end{bmatrix}$, where $\hat{Q}_1, \hat{P}_1 \in \mathbb{R}^{r\times r}$ are nonnegative definite. By Theorem 6.2.5, p. 123 of [55], there exists invertible $\Psi \in \mathbb{R}^{r\times r}$ such that $\tilde{\Lambda}_{\hat{Q}} \triangleq \Psi\hat{Q}_1\Psi^{T}$ and $\tilde{\Lambda}_{\hat{P}} = \Psi^{-T}\hat{P}_1\Psi^{-1}$ are nonnegative diagonal. Because of (3.14), it is clear that Ψ can be chosen so so that $\tilde{\Lambda}_{\hat{Q}} = \begin{bmatrix}\Lambda_{\hat{Q}} & 0\\ 0 & 0\end{bmatrix}$ and $\tilde{\Lambda}_{\hat{P}} = \begin{bmatrix}\Lambda_{\hat{P}} & 0\\ 0 & 0\end{bmatrix}$, where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \mathbb{R}^{n_c\times n_c}$ are positive diagonal. Thus (3.20) holds with $\Phi \triangleq \begin{bmatrix}\Psi & 0\\ 0 & I_{\infty}\end{bmatrix}$. From (3.20) it follows that $\hat{Q}\hat{P} = \Phi^{-1}\begin{bmatrix}\Lambda_Q\Lambda_P & 0\\ 0 & 0\end{bmatrix}\Phi$. Now define $\overline{G} = [I_{n_c} \quad 0]\Phi^{-*}$, $\overline{M} = \Lambda_{\hat{Q}}\Lambda_{\hat{P}}$ and $\overline{\Gamma} = [I_{n_c} \quad 0]\Phi$ so that (3.7)' and (3.8)' are satisfied. By the second part of Proposition 3.2 there exists invertible $S \in \mathbb{R}^{n_c\times n_c}$ such that $G = S^{T}\overline{G}$, $M = S^{-1}\overline{M}S$ and $\Gamma = S^{-1}\overline{\Gamma}$.

Since M and $\overline{M}$ have the same eigenvalues, $\overline{M} = \Lambda$ (modulo an ordering of the diagonal elements) and thus (3.21a) holds. Finally, (3.21b) follows from

$$\tau = G^*\Gamma = \overline{G}^*\overline{\Gamma} = \Phi^{-1} \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} \Phi. \quad \Box$$

Remark 3.1. Proposition 3.3 shows that $\lambda_1, \ldots, \lambda_{n_c}$ are the positive eigenvalues of $\hat{Q}\hat{P}$.

Remark 3.2. The simultaneous diagonalization in (3.20) has been effected by a contragredient transformation ([55,58]). For applications of this type of transformation to model reduction and realization problems see [12, 59-61]. Simultaneous diagonalization of operators is discussed in [53], p. 181.

The following result permits the precise handling of the unbounded operator A in (3.9), (3.17) and (3.18).

Proposition 3.4. The following relations hold:

$$\rho(G) = \rho(\Gamma) = \rho(\tau) = n_c, \qquad (3.23a,b,c)$$

$$\tau: \underline{H} \to \underline{D}(A), \; \tau^*: \underline{H} \to \underline{D}(A^*), \qquad (3.24a,b)$$

$$G^*: \mathbb{R}^{n_c} \to \underline{D}(A), \; \Gamma^*: \mathbb{R}^{n_c} \to \underline{D}(A^*). \qquad (3.25a,b)$$

Proof. From (3.8) and (2.1) it follows that $n_c = \rho(\Gamma G^*) \leq \min\left\{\rho(\Gamma), \rho(G^*)\right\}$. Since $\rho(\Gamma) \leq n_c$, $\rho(G) = \rho(G^*)$ and $\rho(G) \leq n_c$, (3.23a) and (3.23b) hold. To show (3.23c) either note (3.21b) or use (3.14a) and (3.22) to obtain

$$n_c = \rho(\hat{Q}) = \rho(\tau\hat{Q}) \leq \rho(\tau) = \rho(G^*\Gamma) \leq \rho(\Gamma) = n_c.$$

To prove (3.24a) note that (3.22a) implies $\underline{R}(\hat{Q}) \subset \underline{R}(\tau)$ and thus $\rho(\hat{Q}) = \rho(\tau)$ implies $\underline{R}(\hat{Q}) = \underline{R}(\tau)$, and similarly for (3.24b). Finally, (3.25) follows from (3.23), (3.24), the definition $\tau = G^*\Gamma$ and the fact that $\tau^* = \Gamma^*G$. $\Box$

Since the domain of A may not be all of $\underline{H}$, expressions involving A require special interpretation. First note that because of the range condition (3.25a), the expression (3.9) indeed represents an $n_c \times n_c$ matrix (see, e.g., [6], p. 80). Similarly, because of (3.25b), A_c^T is given by

$$A_c^T = G(A^* - \bar{\Sigma}Q - P\Sigma)\Gamma^* . \tag{3.26}$$

With regard to (3.15), note that because of (3.12a), the right-hand side of (3.15) is a linear operator with domain $\underline{D}(A^*)$. Since $\Theta \triangleq -\tau Q\bar{\Sigma}Q - Q\bar{\Sigma}Q\tau^* + V_1 + \tau Q\bar{\Sigma}Q\tau^*$ is continuous on $\underline{D}(A^*)$, $AQ + QA^*$ has a continuous extension on $\underline{H}$ given precisely by $-\Theta$. Similar remarks apply to (3.16). Analogous domain conditions were obtained in [5] for a deterministic infinite-dimensional linear-quadratic control problem with full-state feedback. Finally, because of (3.24) the right-hand sides of (3.17) and (3.18) denote bounded linear operators on all of $\underline{H}$.

It is useful to present an alternative form of the optimal projection equations (3.15)-(3.18). For convenience define the notation

$$\tau_\perp \triangleq I_{\underline{H}} - \tau .$$

<u>Proposition 3.5.</u> Equations (3.15)-(3.18) are equivalent, respectively, to

$$0 = AQ + QA^* + V_1 - Q\bar{\Sigma}Q + \tau_\perp Q\bar{\Sigma}Q\tau_\perp^* , \tag{3.27}$$

$$0 = A^*P + PA + R_1 - P\Sigma P + \tau_\perp^* P\Sigma P\tau_\perp , \tag{3.28}$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^* + Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^* , \tag{3.29}$$

$$0 = (A - Q\bar{\Sigma})^*\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_\perp^* P\Sigma P\tau_\perp . \tag{3.30}$$

Proof. The equivalence of (3.27) and (3.28) to (3.15) and (3.16) is immediate. Using (3.22a) in the form $\hat{Q} = \hat{Q}\tau^*$, we obtain (3.17) = (3.29) τ^*. Conversely, from (3.22a) and $[(A - \Sigma P)\hat{Q}]^* = \hat{Q}(A - \Sigma P)^*$ (see, e.g., [6], p. 80) it follows that (3.29) = (3.17) + (3.17)* - τ(3.17). Similarly, (3.18) and (3.30) are equivalent. □

The form of the optimal projection equations (3.27)-(3.30) helps demonstrate the relationship between the Main Theorem and the classical LQG result when dim $\underline{H} = n < \infty$. In this case we need only note that the (G,M,Γ)-factorization of $\hat{Q}\hat{P}$ in the "full-order" case $n_c = n$ is given by $G = \Gamma = I_n$ and $M = \hat{Q}\hat{P}$. Since $\tau = I_n$, and thus $\tau_1 = 0$, (3.27) and (3.28) reduce to the standard observer and regulator Riccati equations and (3.9)-(3.11) yield the usual LQG expressions. Furthermore, note that in the full-order case

$$A_c = A + BC_c - B_cC \tag{3.31}$$

and (3.29) and (3.31) can be written as

$$0 = (A_c + B_cC)\hat{Q} + \hat{Q}(A_c + B_cC)^T + B_cV_2B_c^T , \tag{3.32}$$

$$0 = (A_c - BC_c)^T\hat{P} + \hat{P}(A_c - BC_c) + C_c^TR_2C_c. \tag{3.33}$$

Since, as is well known, the stability of $\tilde{A}$ corresponds to the stability of $A + BC_c = A_c + B_cC$ and $A - B_cC = A_c - BC_c$, it follows from standard results (e.g., [62], pp. 48, 277) that the positive-definiteness conditions (3.14a,b) are equivalent to the assumption that (A_c,B_c,C_c) is controllable and observable.

To obtain a geometric interpretation of the optimal projection we introduce the quasi-full-state estimate

$$\hat{x}(t) \triangleq G^*x_c(t) \in \underline{H}$$

so that $\tau\hat{x}(t) = \hat{x}(t)$ and $x_c(t) = \Gamma\hat{x}(t)$. Now, the closed-loop system (3.1)-(3.4) can be written as

$$\dot{x}(t) = Ax(t) - B\hat{C}_c\tau\hat{x}(t) + H_1w(t), \tag{3.34}$$

$$\dot{\hat{x}}(t) = \tau(A + B\hat{C}_c - \hat{B}_cC)\tau\hat{x}(t) + \tau\hat{B}_c(Cx(t) + H_2w(t)), \tag{3.35}$$

where (3.35) is interpreted in the sense of (3.34) since $\hat{x}(t) \in \underline{H}$ and where

$$\hat{B}_c \triangleq QC^*V_2^{-1}, \quad \hat{C}_c \triangleq -R_2^{-1}B^*P.$$

It can thus be seen that the geometric structure of the quasi-full-order compensator is entirely dictated by the projection τ. In particular, control inputs $\tau\hat{x}(t)$ determined by (3.35) are contained in $\underline{R}(\tau)$ and sensor inputs $\tau\hat{B}_cy(t)$ are annihilated unless they are contained in $[\underline{N}(\tau)]^{\perp} = \underline{R}(\tau^*)$. Consequently, $\underline{R}(\tau)$ and $\underline{R}(\tau^*)$ are the control and observation subspaces, respectively, of the compensator. Since τ is not necessarily an orthogonal projection, these (finite-dimensional) subspaces may be different.

From the form of (3.35) it is tempting to suggest that the optimal fixed-order dynamic compensator can be obtained by projecting the full-order (infinite-dimensional) LQG compensator. However, this is generally impossible for the following simple reason. Although the expressions for A_c, B_c and C_c in (3.9)-(3.11) have the <u>form</u> of a projection of the full-order LQG compensator, the operators Q and P in (3.9)-(3.11) are <u>not</u> the solutions of the usual LQG Riccati equations but instead must be obtained by simultaneously solving all four coupled equations (3.15)-(3.18). This observation reinforces the statement made in Section 1 that the optimal fixed-order dynamic compensator cannot in general be obtained by LQG followed by closed-loop controller reduction as in [14] and [15].

We now give an explicit characterization of the optimal projection in terms of $\hat{Q}$ and $\hat{P}$. Since $\hat{Q}\hat{P}$ has finite rank, its Drazin inverse $(\hat{Q}\hat{P})^D$ exists (see Theorem 6, p. 108 of [63]) and, since $(\hat{Q}\hat{P})^2 = G^* M^2 \Gamma$, and hence $\rho(\hat{Q}\hat{P})^2 = \rho(\hat{Q}\hat{P})$, the "index" of $\hat{Q}\hat{P}$ (see [63,64]) is 1. In this case the Drazin inverse is traditionally called the group inverse and is denoted by $(\hat{Q}\hat{P})^{\#}$ (see, e.g., [64], p. 124 or [65]).

Proposition 3.6. The optimal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}. \tag{3.36}$$

Proof. It is easy to verify that the conditions characterizing the Drazin inverse ([63]) for the case that $\hat{Q}\hat{P}$ has index 1 are satisfied by $G^* M^{-1}\Gamma$. Hence $(\hat{Q}\hat{P})^{\#} = G^* M^{-1}\Gamma$ and (3.8) implies (3.36). $\square$

We now give an alternative characterization of the optimal projection by introducing the following notation from [51], p. 73. For $\phi, \psi \in \underline{H}$ define the operator $\phi \otimes \psi \in \underline{B}(\underline{H})$ by

$$(\phi \otimes \psi)x \triangleq \langle x, \phi \rangle \psi, \quad x \in \underline{H},$$

and note that $\rho(\phi \otimes \psi) = 1$ if ϕ and ψ are both nonzero and $(\phi \otimes \psi)^* = \psi \otimes \phi$. Using this notation, (3.21a) can be written as

$$\Phi\hat{Q}\hat{P}\Phi^{-1} = \sum_{i=1}^{n_c} \lambda_i \xi_i \otimes \xi_i, \tag{3.37}$$

where $\left\{\xi_i\right\}_{i=1}^{\infty}$ is an orthonormal basis for $\underline{H}$. In terms of the Riesz bases (see e.g., [52], p. 309)

$$\phi_i \triangleq \Phi^* \xi_i, \quad \psi_i \triangleq \Phi^{-1}\xi_i, \quad i = 1,2,\ldots,$$

(3.37) is equivalent to

$$\widehat{\widehat{QP}} = \sum_{i=1}^{n_c} \lambda_i \, \phi_i \otimes \psi_i, \tag{3.38}$$

which can be regarded as a specialized spectral decomposition of a semisimple operator. We emphasize that, in contrast to the singular value decomposition for compact nonnormal operators (see, e.g., [50], p. 261), the λ_i in (3.38) are eigenvalues of $\widehat{\widehat{QP}}$ (see Remark 3.1), not singular values. Moveover, although $\left\{\phi_i\right\}_{i=1}^{\infty}$ and $\left\{\psi_i\right\}_{i=1}^{\infty}$ are bases for $\underline{H}$, they are not necessarily orthogonal. They are, however, biorthonormal, i.e., $<\phi_i, \psi_j> = \delta_{ij}$, and hence $\phi_i \otimes \psi_i$ is a rank-one projection and $(\phi_i \otimes \psi_i)(\phi_j \otimes \psi_j) = 0$, $i \neq j$. Since τ is a rank n_c projection, it is not surprising that τ is given precisely by

$$\tau = \sum_{i=1}^{n_c} \phi_i \otimes \psi_i. \tag{3.39}$$

The following result formalizes the above observations.

Proposition 3.7. There exist biorthonormal linearly independent sets $\left\{\psi_i\right\}_{i=1}^{n_c} \subset \underline{D}(A)$ and $\left\{\phi_i\right\}_{i=1}^{n_c} \subset \underline{D}(A^*)$ such that (3.38) and (3.39) hold .

Furthermore, if the (G,M,Γ)-factorization of $\widehat{\widehat{QP}}$ is chosen such that $M = \Lambda$, then, for all $x \in \underline{H}$,

$$Gx = (<x, \psi_1>, \ldots, <x, \psi_{n_c}>)^T ,$$

$$\Gamma x = (<x, \phi_1>, \ldots, <x, \phi_{n_c}>)^T .$$

Remark 3.3. Note that $\widehat{\widehat{PQ}}$ and τ^* are given by

$$\widehat{\widehat{PQ}} = \sum_{i=1}^{n_c} \lambda_i \, \psi_i \otimes \phi_i,$$

$$\tau^* = \sum_{i=1}^{n_c} \psi_i \otimes \phi_i,$$

and, for all $y \triangleq (y_1,\ldots,y_{n_c})^T \in \mathbb{R}^{n_c}$, G^* and Γ^* satisfy

$$G^* y = \sum_{i=1}^{n_c} y_i \psi_i,$$

$$\Gamma^* y = \sum_{i=1}^{n_c} y_i \phi_i.$$

4. PROOF OF THE MAIN THEOREM

We state and prove a series of lemmas which allow us to compute the Frechet derivatives of J with respect to A_c, B_c and C_c. Requiring that these derivatives vanish leads to the necessary conditions in their "primitive" form. A transformation of variables then leads to the form of the necessary conditions (3.9)-(3.18).

Let "u-lim" denote the uniform limit (i.e., limit in operator norm) for bounded linear operators ([50], p. 150) and, for strongly continuous $S(t) \in \underline{B}(\underline{H})$, $t \geq 0$, interpret the strong integral $\int_{t_1}^{t_2} S(t)\,dt$ according to $\int_{t_1}^{t_2} S(t)z\,dt$, $z \in \underline{H}$ ([50], p. 152). Also recall the standard fact ([6], p. 186) that $(e^{At})^* = e^{A^*t}$ and similarly for $\widetilde{A}$. Throughout this section let $(A_c, B_c, C_c) \in \underline{A}_+$ and let $M, \beta > 0$ satisfy (3.6).

To begin, note that the closed-loop system (3.1)-(3.4) can be written as

$$\dot{\widetilde{x}}(t) = \widetilde{A}\widetilde{x}(t) + \widetilde{H}w(t), \tag{4.1}$$

where

$$\widetilde{H} \triangleq \begin{bmatrix} H_1 \\ B_c H_2 \end{bmatrix} \in \underline{B}_2(\underline{H}' \oplus \mathbb{R}^{\ell}).$$

For convenience define the nonnegative-definite operator

$$\widetilde{V} \triangleq \widetilde{H}\widetilde{H}^* = \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^T \end{bmatrix} \in \underline{B}_1(\widetilde{\underline{H}}).$$

In terms of the augmented state $\widetilde{x}(t)$, the performance criterion (3.5) becomes

$$J(A_c, B_c, C_c) = \lim_{t\to\infty} \mathbb{E}\langle \widetilde{R}\widetilde{x}(t), \widetilde{x}(t)\rangle, \tag{4.2}$$

where the nonnegative-definite operator $\widetilde{R}$ is defined by

$$\widetilde{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^T R_2 C_c \end{bmatrix} \in \underline{B}_1(\widetilde{\underline{H}}).$$

To write (4.2) in terms of the covariance of $\widetilde{x}(t)$, recall ([6], p. 308) that the covariance "$\mathbb{E}[(\xi - \mathbb{E}\xi)(\xi - \mathbb{E}\xi)^*]$" of a Hilbert-space-valued weak random variable ξ is defined to be the nonnegative-definite operator S which satisfies

$$\langle Sy,z\rangle = \mathbb{E}\langle \xi - \mathbb{E}\xi, y\rangle\langle \xi - \mathbb{E}\xi, z\rangle$$

for all y,z in the Hilbert space. Hence define ([6]), p. 317)

$$\widetilde{Q}(t) \triangleq \mathbb{E}[(\widetilde{x}(t) - \mathbb{E}\widetilde{x}(t))(\widetilde{x}(t) - \mathbb{E}\widetilde{x}(t))^*].$$

<u>Lemma 4.1</u>. $\widetilde{Q} \triangleq \underset{t\to\infty}{\text{u-lim}}\, \widetilde{Q}(t)$ exists and is given by

$$\widetilde{Q} = \int_0^\infty e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^* t}dt. \tag{4.3}$$

Furthermore,

$$J(A_c,B_c,C_c) = \operatorname{tr} \widetilde{Q}\widetilde{R}. \tag{4.4}$$

<u>Proof</u>. First compute (as in [6], p. 317)

$$\begin{aligned}
\langle \widetilde{Q}(t)\widetilde{y},\widetilde{z}\rangle &= \mathbb{E}\langle \widetilde{x}(t) - e^{\widetilde{A}t}\mathbb{E}\widetilde{x}(0),\widetilde{y}\rangle\langle \widetilde{x}(t) - e^{\widetilde{A}t}\mathbb{E}\widetilde{x}(0),\widetilde{z}\rangle \\
&= \mathbb{E}\langle \int_0^t e^{\widetilde{A}(t-s)}\widetilde{H}\widetilde{w}(s)ds,\widetilde{y}\rangle\langle \int_0^t e^{\widetilde{A}(t-\sigma)}\widetilde{H}\widetilde{w}(\sigma)d\sigma,\widetilde{z}\rangle + \langle \widetilde{Q}(0)e^{\widetilde{A}^* t}\widetilde{y}, e^{\widetilde{A}^* t}\widetilde{z}\rangle \\
&= \mathbb{E}\int_0^t\int_0^t \langle \widetilde{w}(s),\widetilde{H}^* e^{\widetilde{A}^*(t-s)}\widetilde{y}\rangle\langle \widetilde{w}(\sigma),\widetilde{H}^* e^{\widetilde{A}^*(t-\sigma)}\widetilde{z}\rangle\, ds\, d\sigma + \langle e^{\widetilde{A}t}\widetilde{Q}(0)e^{\widetilde{A}^* t}\widetilde{y},\widetilde{z}\rangle \\
&= \int_0^t \langle e^{\widetilde{A}(t-s)}\widetilde{V}e^{\widetilde{A}^*(t-s)}\widetilde{y},\widetilde{z}\rangle\, ds + \langle e^{\widetilde{A}t}\widetilde{Q}(0)e^{\widetilde{A}^* t}\widetilde{y},\widetilde{z}\rangle,
\end{aligned}$$

which shows that $\tilde{Q}(t)$ is given by

$$\tilde{Q}(t) = e^{\tilde{A}t}\tilde{Q}(0)e^{\tilde{A}^* t} + \int_0^t e^{\tilde{A}s}\tilde{V}e^{\tilde{A}^* s}ds.$$

Clearly, (4.3) makes sense as a strong integral since

$$\begin{aligned} ||\tilde{Q}|| &\leq \int_0^\infty ||e^{\tilde{A}t}\tilde{V}e^{\tilde{A}^* t}||dt \\ &\leq M^2||\tilde{V}|| \int_0^\infty e^{-2\beta t}dt \\ &< \infty . \end{aligned}$$

To demonstrate uniform convergence it need only be noted that

$$\begin{aligned} ||\tilde{Q} - \tilde{Q}(t)|| &= \sup_{||\tilde{y}||=1} ||(\tilde{Q} - \tilde{Q}(t))\tilde{y}|| \\ &= \sup_{||\tilde{y}||=1} ||\int_t^\infty e^{\tilde{A}s}\tilde{V}e^{\tilde{A}^* s}\tilde{y}\, ds - e^{\tilde{A}t}\tilde{Q}(0)e^{\tilde{A}^* t}\tilde{y}|| \\ &\leq \int_t^\infty ||e^{\tilde{A}s}\tilde{V}e^{\tilde{A}^* s}||ds + ||e^{\tilde{A}t}\tilde{Q}(0)e^{\tilde{A}^* t}|| \\ &\leq \frac{1}{2} M^2||\tilde{V}|| \beta^{-1}e^{-2\beta t} + ||\tilde{Q}(0)||e^{-2\beta t}. \end{aligned}$$

Next, let $\{\phi_i\}_{i=1}^\infty$ be an orthonormal basis for $\underline{\tilde{H}}$ and use Parseval's equality to obtain

$$\begin{aligned} J(A_c,B_c,C_c) &= \lim_{t\to\infty} \mathbb{E}|| \tilde{R}^{1/2}\tilde{x}(t)||^2 \\ &= \lim_{t\to\infty} \mathbb{E}\sum_{i=1}^\infty <\tilde{R}^{1/2}\tilde{x}(t), \phi_i>^2. \end{aligned}$$

Since $f_n(t) \triangleq \sum_{i=1}^n <\tilde{R}^{1/2}\tilde{x}(t), \phi_i>^2$, $t\geq 0$, is nonnegative for each n and is

increasing in n for each t with limit $<\widetilde{R}\widetilde{x}(t),\widetilde{x}(t)>$, monotone convergence permits expectation-limit interchange. Hence using $\mathbb{E}\widetilde{x}(t) = e^{\widetilde{A}t}\mathbb{E}\widetilde{x}(0)$ we have

$$J(A_c,B_c,C_c) = \lim_{t\to\infty} \sum_{i=1}^{\infty} \mathbb{E}<\widetilde{x}(t),\widetilde{R}^{1/2}\phi_i>^2$$

$$= \lim_{t\to\infty} \sum_{i=1}^{\infty} [<\widetilde{Q}(t)\widetilde{R}^{1/2}\phi_i,\widetilde{R}^{1/2}\phi_i> + <e^{\widetilde{A}t}\mathbb{E}\widetilde{x}(0),\widetilde{R}^{1/2}\phi_i>^2]$$

$$= \lim_{t\to\infty} \left\{ tr[\widetilde{R}^{1/2}\widetilde{Q}(t)\widetilde{R}^{1/2}] + ||\widetilde{R}^{1/2}e^{\widetilde{A}t}\mathbb{E}\widetilde{x}(0)||^2 \right\}$$

which by Corollary 2.1 yields (4.4). $\square$

We shall also require the "dual" of $\widetilde{Q}$ given by

$$\widetilde{P} = \int_0^\infty e^{\widetilde{A}^* t}\widetilde{R}e^{\widetilde{A}t}dt. \tag{4.5}$$

Since $\widetilde{V}$ and $\widetilde{R}$ are nonnegative definite it is readily seen that $\widetilde{Q}$ and $\widetilde{P}$ are also nonnegative definite.

<u>Lemma 4.2</u>. $\widetilde{Q},\widetilde{P} \in \underline{B}_1(\underline{\widetilde{H}})$.

<u>Proof</u>. It suffices to consider $\widetilde{Q}$ only since the situation for $\widetilde{P}$ is exactly analogous. Since $\widetilde{Q}$ is nonnegative definite, Lemma 2.3 can be used. Letting $\left\{\phi_i\right\}_{i=1}^{\infty}$ be an orthonormal basis for $\underline{\widetilde{H}}$, we have

$$tr\ \widetilde{Q} = \sum_{i=1}^{\infty} <\widetilde{Q}\phi_i,\ \phi_i>$$

$$= \sum_{i=1}^{\infty} <\int_0^\infty e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^* t}\phi_i dt,\ \phi_i>$$

$$= \lim_{n\to\infty} \int_0^\infty \sum_{i=1}^{n} <\widetilde{V}e^{\widetilde{A}^* t}\phi_i, e^{\widetilde{A}^* t}\phi_i>\ dt.$$

Let $f_n(t)$ denote the above integrand. Since $\widetilde{V}$ is nonnegative definite, $\{f_n(\cdot)\}$ is a monotonically increasing sequence of nonnegative functions such that $f_n(t) \to \operatorname{tr} e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}$, $t \geq 0$. Hence, by monotone convergence and Lemma 2.2,

$$\begin{aligned}
\operatorname{tr} \widetilde{Q} &= \int_0^\infty \operatorname{tr}[e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}]dt \\
&= \int_0^\infty ||[e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}||_1 dt \\
&\leq M^2 ||\widetilde{V}||_1 \int_0^\infty e^{-2\beta t}dt \\
&< \infty . \quad \square
\end{aligned}$$

Lemma 4.3. *With $\widetilde{Q}$ and $\widetilde{P}$ given by (4.3) and (4.5) it follows that*

$$\operatorname{tr} \widetilde{Q}\widetilde{R} = \operatorname{tr} \widetilde{V}\widetilde{P}. \tag{4.6}$$

Proof. For any orthonormal basis $\{\phi_i\}_{i=1}^\infty$ of $\underline{\widetilde{H}}$ we have

$$\begin{aligned}
\operatorname{tr} \widetilde{Q}\widetilde{R} &= \operatorname{tr} \widetilde{R}\widetilde{Q} \\
&= \sum_{i=1}^\infty \langle \widetilde{R}\int_0^\infty e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}\phi_i dt, \ \phi_i\rangle \\
&= \lim_{n\to\infty} \int_0^\infty \sum_{i=1}^n \langle \widetilde{R}e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}\phi_i, \ \phi_i\rangle dt.
\end{aligned}$$

Letting $f_n(t)$ denote the above integrand it follows that $f_n(t) \to \operatorname{tr} \widetilde{R}e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}$, $t \geq 0$, and

$$\begin{aligned}
|f_n(t)| &\leq \sum_{i=1}^\infty |\langle e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t}\phi_i, \widetilde{R}\phi_i| \\
&\leq M^2 ||\widetilde{V}|| e^{-2\beta t} \sum_{i=1}^\infty ||\widetilde{R}\phi_i||.
\end{aligned}$$

If $\{\phi_i\}_{i=1}^{\infty}$ is chosen to be the set of orthonormal eigenvectors of $\widetilde{R}$ then Lemma 2.1 implies $\sum_{i=1}^{\infty} ||\widetilde{R}\phi_i|| = ||\widetilde{R}||_1$ and thus $|f_n(t)|$ is bounded on $[0,\infty)$ by an integrable function. Hence by dominated convergence,

$$\begin{aligned} \operatorname{tr} \widetilde{Q}\widetilde{R} &= \int_0^{\infty} \operatorname{tr}[\widetilde{R}e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^* t}]\, dt \\ &= \int_0^{\infty} \operatorname{tr}[e^{\widetilde{A}^* t}\widetilde{R}e^{\widetilde{A}t}\widetilde{V}]\, dt \\ &= \int_0^{\infty} \sum_{i=1}^{\infty} \langle \widetilde{V}\phi_i, e^{\widetilde{A}^* t}\widetilde{R}e^{\widetilde{A}t} \rangle\, dt. \end{aligned}$$

And again using dominated convergence,

$$\begin{aligned} \operatorname{tr} \widetilde{Q}\widetilde{R} &= \sum_{i=1}^{\infty} \int_0^{\infty} \langle \widetilde{V}\phi_i, e^{\widetilde{A}^* t}\widetilde{R}e^{\widetilde{A}t}\phi_i \rangle\, dt \\ &= \sum_{i=1}^{\infty} \langle \widetilde{V}\phi_i, \int_0^{\infty} e^{\widetilde{A}^* t}\widetilde{R}e^{\widetilde{A}t}\phi_i\, dt \rangle \\ &= \operatorname{tr} \widetilde{V}\widetilde{P}. \quad \square \end{aligned}$$

The next result is important in that it allows us to treat $\widetilde{Q}$ and $\widetilde{P}$ as solutions of dual algebraic Lyapunov equations. For a similar result involving groups rather than semigroups see [50], pp. 555-557.

<u>Lemma 4.4</u>. $\widetilde{Q}$ is given by (4.3) if and only if $\widetilde{Q} \in \underline{B}(\underline{\widetilde{H}})$ satisfies

$$\widetilde{Q}: \ \underline{D}(\widetilde{A}^*) \to \underline{D}(\widetilde{A}), \tag{4.7}$$

$$0 = \widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^* + \widetilde{V}, \tag{4.8}$$

where (4.8) holds in the sense discussed in Section 3. Furthermore, $\widetilde{P}$ is given by (4.5) if and only if $\widetilde{P} \in \underline{B}(\underline{\widetilde{H}})$ satisfies

$$\widetilde{P}: \ \underline{D}(\widetilde{A}) \to \underline{D}(\widetilde{A}^*), \tag{4.9}$$

$$0 = \widetilde{A}^*\widetilde{P} + \widetilde{P}\widetilde{A} + \widetilde{R}. \tag{4.10}$$

Proof. We consider $\widetilde{Q}$ only. To prove necessity let $t' > 0$. Then for all $t \in [0,t')$ and $\widetilde{x} \in \underline{D}(\widetilde{A}^*)$ we can write

$$e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^*t'}\widetilde{x} = \int_0^\infty e^{\widetilde{A}(t+s)}\widetilde{V}e^{\widetilde{A}^*(t'+s)}\widetilde{x}\, ds$$

$$= \int_t^\infty e^{\widetilde{A}\sigma}\widetilde{V}e^{\widetilde{A}^*\sigma}e^{\widetilde{A}^*(t'-t)}\widetilde{x}\, d\sigma.$$

Hence,

$$\frac{d}{dt} e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^*t'}\widetilde{x} = -\int_t^\infty e^{\widetilde{A}\sigma}\widetilde{V}e^{\widetilde{A}^*\sigma}e^{\widetilde{A}^*(t'-t)}\widetilde{A}^*\widetilde{x}\, d\sigma - e^{\widetilde{A}t}\widetilde{V}e^{\widetilde{A}^*t'}\widetilde{x}, \tag{4.11}$$

which shows that $e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^*t'}$ is strongly differentiable with respect to t for all $t \in [0,t')$. In particular, setting $t = 0$ it follows that $\widetilde{Q}e^{\widetilde{A}^*t'}\widetilde{x} \in \underline{D}(\widetilde{A})$ for all $\widetilde{x} \in \underline{D}(\widetilde{A}^*)$ (see, e.g., [6], p. 173, or [50], p. 485). Performing the differentiation on the left-hand side of (4.11) and setting $t = 0$ yields

$$\widetilde{A}\widetilde{Q}e^{\widetilde{A}^*t'}\widetilde{x} = -\int_0^\infty e^{\widetilde{A}\sigma}\widetilde{V}e^{\widetilde{A}^*\sigma}e^{\widetilde{A}^*t'}\widetilde{A}^*\widetilde{x}\, d\sigma - \widetilde{V}e^{\widetilde{A}^*t'}\widetilde{x}. \tag{4.12}$$

Now fix $\widetilde{x} \in \underline{D}(\widetilde{A}^*)$. Then for $\left\{t_i\right\}_{i=1}^\infty$, $t_i > 0$, $t_i \to 0$, we have

$$\widetilde{Q}e^{\widetilde{A}^*t_i}\widetilde{x} \in \underline{D}(\widetilde{A}),\quad i = 1,2,3,\ldots,$$

$$\widetilde{Q}e^{\widetilde{A}^*t_i}\widetilde{x} \underset{i\to\infty}{\longrightarrow} \widetilde{Q}\widetilde{x}.$$

Now consider the sequence $\left\{\widetilde{A}\widetilde{Q}e^{\widetilde{A}^*t_i}\widetilde{x}\right\}_{i=1}^\infty$. Letting $t'=t_1$ in (4.12) and using dominated convergence to interchange limit and integration ($\widetilde{A}^*\widetilde{x}$ is a fixed element of $\underline{\widetilde{H}}$), it follows that

$$\lim_{i\to\infty} \widetilde{A}\widetilde{Q}e^{\widetilde{A}^* t_i}\widetilde{x} = -\int_0^{\infty} e^{\widetilde{A}\sigma}\widetilde{V}e^{\widetilde{A}^*\sigma}\widetilde{A}^*\widetilde{x}\, d\sigma - \widetilde{V}\widetilde{x}. \tag{4.13}$$

Since $\widetilde{A}$ is closed, $\widetilde{Q}\widetilde{x} \in \underline{D}(\widetilde{A})$. This proves (4.7). Also, since $\widetilde{A}$ is closed we have

$$\lim_{i\to\infty} \widetilde{A}\widetilde{Q}e^{\widetilde{A}^* t_i}\widetilde{x} = \widetilde{A}\widetilde{Q}\widetilde{x},$$

which with (4.13) implies

$$\widetilde{A}\widetilde{Q}\widetilde{x} = -\widetilde{Q}\widetilde{A}^*\widetilde{x} - \widetilde{V}\widetilde{x},$$

and hence

$$(\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^* + \widetilde{V})\widetilde{x} = 0,\ \widetilde{x} \in \underline{D}(\widetilde{A}^*),$$

as desired.

To prove sufficiency let $\widetilde{x} \in \underline{D}(\widetilde{A})$. Then $e^{\widetilde{A}^* t}\widetilde{x} \in \underline{D}(\widetilde{A}^*)$, $t \geq 0$, and hence

$$\frac{d}{dt}\, e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^* t}\widetilde{x} = e^{\widetilde{A}t}(\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^*)e^{\widetilde{A}^* t}\widetilde{x}.$$

Thus

$$e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^* t}\widetilde{x} - \widetilde{Q}\widetilde{x} = \int_0^t e^{\widetilde{A}s}(\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^*)e^{\widetilde{A}^* s}\widetilde{x}\, ds,\ \widetilde{x} \in \underline{D}(\widetilde{A}^*).$$

Extending $\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^*$ to all of $\underline{\widetilde{H}}$ we obtain

$$e^{\widetilde{A}t}\widetilde{Q}e^{\widetilde{A}^* t}\widetilde{x} - \widetilde{Q}\widetilde{x} = -\int_0^t e^{\widetilde{A}s}\widetilde{V}e^{\widetilde{A}^* s}\widetilde{x}\, ds,\ \widetilde{x} \in \underline{\widetilde{H}}.$$

Letting $t \to \infty$ yields (4.3). □

We now introduce some notation which will prove to be most convenient in the following results. For $(A_c', B_c', C_c') \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$ define

$$\delta_{A_c} \triangleq A_c' - A_c, \; \delta_{B_c} \triangleq B_c' - B_c, \; \delta_{C_c} \triangleq C_c' - C_c$$

and

$$||(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})|| \triangleq ||\delta_{A_c}|| + ||\delta_{B_c}|| + ||\delta_{C_c}||.$$

Furthermore, let $\tilde{A}'$, $\tilde{V}'$ and $\tilde{R}'$ denote $\tilde{A}$, $\tilde{V}$ and $\tilde{R}$ with (A_c, B_c, C_c) replaced by (A_c', B_c', C_c') and define

$$\delta_{\tilde{A}} \triangleq \tilde{A}' - \tilde{A} = \begin{bmatrix} 0 & B\delta_{C_c} \\ \delta_{B_c} C & \delta_{A_c} \end{bmatrix},$$

$$\delta_{\tilde{V}} \triangleq \tilde{V}' - \tilde{V} = \begin{bmatrix} 0 & 0 \\ 0 & B_c V_2 \delta_{B_c}^T + \delta_{B_c} V_2 B_c^T + \delta_{B_c} V_2 \delta_{B_c}^T \end{bmatrix},$$

$$\delta_{\tilde{R}} \triangleq \tilde{R}' - \tilde{R} = \begin{bmatrix} 0 & 0 \\ 0 & C_c^T R_2 \delta_{C_c} + \delta_{C_c}^T R_2 C_c + \delta_{C_c}^T R_2 \delta_{C_c} \end{bmatrix}.$$

We shall also write $\widetilde{Q}'$, $\widetilde{P}'$ for $\widetilde{Q}$, $\widetilde{P}$ as given by (4.3) and (4.5) with $\widetilde{A}$, $\widetilde{V}$, $\widetilde{R}$ replaced by $\widetilde{A}'$, $\widetilde{V}'$, $\widetilde{R}'$ and define

$$\delta_{\widetilde{Q}} \triangleq \widetilde{Q}' - \widetilde{Q}, \quad \delta_{\widetilde{P}} \triangleq \widetilde{P}' - \widetilde{P}.$$

<u>Lemma 4.5.</u> $\underline{A}$ is open.

<u>Proof.</u> Let $(A_c, B_c, C_c) \in \underline{A}$ be arbitrary and consider the open set

$$N \triangleq \left\{ (A_c', B_c', C_c') \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} : \right. \tag{4.14}$$

$$\left. \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\| < \beta / 2M\gamma \right\},$$

where $\gamma \triangleq \max\left\{1, \|B\|, \|C\|\right\}$. Then, since $\widetilde{A}' = \widetilde{A} + \delta_{\widetilde{A}}$ and $\delta_{\widetilde{A}} \in \underline{B}(\underline{\widetilde{H}})$ it follows from Theorem 2.1, p. 497 of [50], that for all $(A_c', B_c', C_c') \in N$ and $t \geq 0$,

$$\|e^{\widetilde{A}' t}\| \leq M e^{(-\beta + M\|\delta_{\widetilde{A}}\|)t}$$

$$\leq M e^{\frac{-\beta}{2} t} .$$

Hence, $N \subset \underline{A}$, as desired. $\square$

Lemma 4.6. There exists $c > 0$ such that

$$||\delta_{\tilde{Q}}|| \leq c||(\delta_{A_c},\delta_{B_c},\delta_{C_c})||, \tag{4.15}$$

$$||\delta_{\tilde{P}}|| \leq c||(\delta_{A_c},\delta_{B_c},\delta_{C_c})||, \tag{4.16}$$

for all $(A_c',B_c',C_c') \in N$, where $N \subset \underline{A}$ is the open neighborhood of (A_c,B_c,C_c) defined by (4.14).

Proof. We consider (4.15) only. Since $||e^{\tilde{A}'t}|| \leq Me^{\frac{-\beta}{2}t}$, $t \geq 0$, $(A_c',B_c',C_c') \in N$, it follows that

$$||\delta_{\tilde{Q}}|| \leq \int_0^\infty ||e^{\tilde{A}'t}\tilde{V}'e^{\tilde{A}'^*t} - e^{\tilde{A}t}\tilde{V}e^{\tilde{A}^*t}||\,dt$$

$$\leq \int_0^\infty \Big\{||e^{\tilde{A}'t}||\;||\tilde{V}'||\;||e^{\tilde{A}'^*t} - e^{\tilde{A}^*t}||$$

$$+\,||e^{\tilde{A}'t}||\;||\delta_{\tilde{V}}||\;||e^{\tilde{A}^*t}||$$

$$+\,||e^{\tilde{A}'t} - e^{\tilde{A}t}||\;||\tilde{V}||\;||e^{\tilde{A}^*t}||\Big\}\,dt$$

$$\leq M(||\widetilde{V}|| + ||\delta_{\widetilde{V}}||)\int_0^\infty ||e^{(\widetilde{A}^* + \delta_{\widetilde{A}}^*)t} - e^{\widetilde{A}^* t}||e^{\frac{-\beta}{2}t}\, dt$$

$$+ M^2||\delta_{\widetilde{V}}||\int_0^\infty e^{\frac{-3\beta}{2}t}\, dt$$

$$+ M||\widetilde{V}||\int_0^\infty ||e^{(\widetilde{A}+\widetilde{\delta}_{\widetilde{A}})t} - e^{\widetilde{A}t}||e^{\frac{-\beta}{2}t}\, dt$$

$$= M(2||\widetilde{V}|| + ||\delta_{\widetilde{V}}||)\int_0^\infty ||e^{(\widetilde{A}+\delta_{\widetilde{A}})t} - e^{\widetilde{A}t}||e^{\frac{-\beta}{2}t}\, dt \;+\; \frac{2M^2}{3\beta}||\delta_{\widetilde{V}}||. \qquad (4.17)$$

From [50], p. 497, it follows that the perturbed semigroup $e^{(\widetilde{A}+\delta_{\widetilde{A}})t}$ has an expansion

$$e^{(\widetilde{A}+\delta_{\widetilde{A}})t} = e^{\widetilde{A}t} + \sum_{i=1}^{\infty} U_i(t),\ t \geq 0,$$

where $U_i(t) \in \underline{B}(\underline{H})$, $t \geq 0$, satisfy the estimates

$$||U_i(t)|| \leq M^{i+1}||\delta_{\widetilde{A}}||^i e^{-\beta t} t^i/i! \quad .$$

Hence, for all $(A_c', B_c', C_c') \in N$,

$$||e^{(\widetilde{A}+\delta_{\widetilde{A}})t} - e^{\widetilde{A}t}|| \leq \sum_{i=1}^{\infty} ||U_i(t)||$$

$$\leq Me^{-\beta t}[e^{M||\delta_{\widetilde{A}}||t} - 1]. \qquad (4.18)$$

From (4.17), (4.18) and the relations $||\delta_{\widetilde{A}}|| \le \gamma||(\delta_{A_c},\delta_{B_c},\delta_{C_c})|| < \beta/2M$ and

$$\int_0^\infty \left[e^{M||\delta_{\widetilde{A}}||t} - 1 \right] e^{\frac{-3\beta}{2}t} \, dt < \frac{M\gamma}{3\beta^2}||(\delta_{A_c},\delta_{B_c},\delta_{C_c})||$$

it follows that

$$||\delta_{\widetilde{Q}}|| \le \frac{2M^3\gamma}{3\beta^2}(2||\widetilde{V}|| + ||\delta_{\widetilde{V}}||)||(\delta_{A_c},\delta_{B_c},\delta_{C_c})||$$

$$+ \frac{2M^2}{3\beta}(2||B_cV_2|| \; ||\delta_{B_c}|| + ||V_2|| \; ||\delta_{B_c}||^2),$$

which yields (4.15). $\square$

Since $\widetilde{Q}$, $\widetilde{P} \in \underline{B}(\underline{\widetilde{H}})$ we can write

$$\widetilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^* & Q_2 \end{bmatrix}, \; \widetilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^* & P_2 \end{bmatrix},$$

where $Q_1 \in \underline{B}(\underline{H})$, $Q_{12} \in \underline{B}(\mathbb{R}^{n_c}, \underline{H})$, $Q_2 \in \mathbb{R}^{n_c \times n_c}$ and similarly for P_1, P_{12} and P_2. Note that Q_1, Q_2, P_1 and P_2 are nonnegative definite. Also, define the notation

$$\widetilde{PQ} = \begin{bmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{bmatrix},$$

where

$$Z_1 \triangleq P_1Q_1 + P_{12}Q_{12}^{*}, \qquad Z_{12} \triangleq P_1Q_{12} + P_{12}Q_2,$$

$$Z_{21} \triangleq P_{12}^{*}Q_1 + P_2Q_{12}^{*}, \qquad Z_2 \triangleq P_{12}^{*}Q_{12} + P_2Q_2,$$

and, for $(A_c', B_c', C_c') \in \underline{A}$, let

$$\delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \triangleq J(A_c', B_c', C_c') - J(A_c, B_c, C_c).$$

<u>Lemma 4.7</u>. Let $(A_c', B_c', C_c') \in \underline{A}$. Then

$$\delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) = \underline{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) + o(||(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})||), \tag{4.19}$$

where

$$\underline{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \triangleq 2tr[Z_2\delta_{A_c}] + 2tr[(V_2B_c^TP_2 + CZ_{21}^{*})\delta_{B_c}] \tag{4.20}$$

$$+ 2tr[Q_2C_c^TR_2 + Z_{12}^{*}B)\delta_{C_c}]$$

and

$$\lim_{(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \to 0} ||(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})||^{-1} o(||(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})||) = 0. \tag{4.21}$$

<u>Proof</u>. Combining (4.8) and (4.10) with (4.6), J can be written as

$$J(A_c,B_c,C_c) = tr[\widetilde{Q}\widetilde{R} + \widetilde{P}\widetilde{V}] + \tfrac{1}{2}tr[\widetilde{Q}c\ell(\widetilde{A}^*\widetilde{P} + \widetilde{P}\widetilde{A}) + \widetilde{P}c\ell(\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^*)],$$

and likewise for (A_c',B_c',C_c'), where "$c\ell$" denotes closure (i.e., extension) of a bounded operator to all of $\widetilde{\underline{H}}$. Now using the identity

$$tr[\widetilde{Q}'\widetilde{R}' + \widetilde{P}'\widetilde{V}'] - tr[\widetilde{Q}\widetilde{R} + \widetilde{P}\widetilde{V}] = tr[\widetilde{Q}\delta_{\widetilde{R}} + \widetilde{P}\delta_{\widetilde{V}}] + tr[\delta_{\widetilde{Q}}\widetilde{R}' + \delta_{\widetilde{P}}\widetilde{V}']$$

we can compute

$$\delta_J(\delta_{A_c},\delta_{B_c},\delta_{C_c}) = tr[\widetilde{Q}\delta_{\widetilde{R}} + \widetilde{P}\delta_{\widetilde{V}}]$$

$$+ \tfrac{1}{2}tr[\widetilde{Q}c\ell(\widetilde{A}^*(\widetilde{P}+\delta_{\widetilde{P}}) + (\widetilde{P}+\delta_{\widetilde{P}})\widetilde{A}')] + \tfrac{1}{2}tr[\delta_{\widetilde{Q}}c\ell(\widetilde{A}'^*\widetilde{P}'+\widetilde{P}'\widetilde{A}')]$$

$$+ \tfrac{1}{2}tr[\widetilde{P}c\ell(\widetilde{A}'(\widetilde{Q}+\delta_{\widetilde{Q}}) + (\widetilde{Q}+\delta_{\widetilde{Q}})\widetilde{A}'^*)] + \tfrac{1}{2}tr[\delta_{\widetilde{P}}c\ell(\widetilde{A}'\widetilde{Q}'+\widetilde{Q}'\widetilde{A}'^*)]$$

$$- \tfrac{1}{2}tr[\widetilde{Q}c\ell(\widetilde{A}^*\widetilde{P}+\widetilde{P}\widetilde{A}) + \widetilde{P}c\ell(\widetilde{A}\widetilde{Q}+\widetilde{Q}\widetilde{A}^*)] + tr[\delta_{\widetilde{Q}}\widetilde{R}' + \delta_{\widetilde{P}}\widetilde{V}'].$$

Using $\widetilde{A}' = \widetilde{A}+\delta_{\widetilde{A}}$ and combining the second, fourth and sixth terms yields

$$\delta_J(\delta_{A_c},\delta_{B_c},\delta_{C_c}) = \Lambda + \Omega ,$$

where

$$\Lambda \triangleq tr[\widetilde{Q}\delta_{\widetilde{R}} + \widetilde{P}\delta_{\widetilde{V}}] + \tfrac{1}{2}tr[\widetilde{Q}(\delta_{\widetilde{A}}^*\widetilde{P}+\widetilde{P}\delta_{\widetilde{A}}) + \widetilde{P}(\delta_{\widetilde{A}}\widetilde{Q}+\widetilde{Q}\delta_{\widetilde{A}}^*)]$$

$$= tr[\widetilde{Q}\delta_{\widetilde{R}}+\widetilde{P}\delta_{\widetilde{V}}] + 2tr[\delta_{\widetilde{A}}\widetilde{Q}\widetilde{P}]$$

and

$$\Omega \triangleq \tfrac{1}{2}\mathrm{tr}[\widetilde{Q}c\ell(\widetilde{A}'^{*}\delta_{\widetilde{P}}+\delta_{\widetilde{P}}\widetilde{A}') + \widetilde{P}c\ell(\widetilde{A}'\delta_{\widetilde{Q}}+\delta_{\widetilde{Q}}\widetilde{A}'^{*})]$$

$$+ \tfrac{1}{2}\mathrm{tr}[\delta_{\widetilde{Q}}c\ell(\widetilde{A}'^{*}\widetilde{P}'+\widetilde{P}'\widetilde{A}') + \delta_{\widetilde{P}}c\ell(\widetilde{A}'\widetilde{Q}'+\widetilde{Q}'\widetilde{A}'^{*})]$$

$$+ \mathrm{tr}[\delta_{\widetilde{Q}}\widetilde{R}' + \delta_{\widetilde{P}}\widetilde{V}'].$$

Computing

$$\mathrm{tr}[\widetilde{Q}\delta_{\widetilde{R}} + \widetilde{P}\delta_{\widetilde{V}}] = 2\mathrm{tr}[V_2B_c^TP_2\delta_{B_c}] + 2\mathrm{tr}[Q_2C_c^TR_2\delta_{C_c}]$$

$$+ \mathrm{tr}[P_2\delta_{B_c}V_2\delta_{B_c}^T + Q_2\delta_{C_c}^TR_2\delta_{C_c}]$$

and

$$2\mathrm{tr}[\delta_{\widetilde{A}}\widetilde{Q}\widetilde{P}] = 2\mathrm{tr}[Z_2\delta_{A_c}] + 2\mathrm{tr}[CZ_{21}^{*}\delta_{B_c}] + 2\mathrm{tr}[Z_{12}^{*}B\delta_{C_c}]$$

and retaining first-order terms, we obtain (4.20).

To evaluate Ω, use (4.8) and (4.10) to replace $\widetilde{R}'$ and $\widetilde{V}'$ in the last term in Ω and write $\widetilde{A}' = \widetilde{A}+\delta_{\widetilde{A}}$, to obtain

$$\Omega = \tfrac{1}{2}\mathrm{tr}[\widetilde{Q}c\ell(\widetilde{A}^{*}\delta_{\widetilde{P}}+\delta_{\widetilde{P}}\widetilde{A}) + \widetilde{P}c\ell(\widetilde{A}\delta_{\widetilde{Q}}+\delta_{\widetilde{Q}}\widetilde{A}^{*})]$$

$$+ \tfrac{1}{2}\mathrm{tr}[\widetilde{Q}(\delta_{\widetilde{A}}^{*}\delta_{\widetilde{P}}+\delta_{\widetilde{P}}\delta_{\widetilde{A}}) + \widetilde{P}(\delta_{\widetilde{A}}\delta_{\widetilde{Q}}+\delta_{\widetilde{Q}}\delta_{\widetilde{A}}^{*})] \qquad (4.22)$$

$$- \tfrac{1}{2}\mathrm{tr}[\delta_{\widetilde{Q}}c\ell(\widetilde{A}'^{*}\widetilde{P}'+\widetilde{P}'\widetilde{A}') + \delta_{\widetilde{P}}c\ell(\widetilde{A}'\widetilde{Q}'+\widetilde{Q}'\widetilde{A}'^{*})].$$

Next we note that

$$\mathrm{tr}[\widetilde{Q}c\ell(\widetilde{A}^{*}\delta_{\widetilde{P}} + \delta_{\widetilde{P}}\widetilde{A}^{*})] = \mathrm{tr}[\delta_{\widetilde{P}}c\ell(\widetilde{A}\widetilde{Q}+\widetilde{Q}\widetilde{A}^{*})]. \qquad (4.23)$$

To see this we observe that by arguments similar to those used in the proof of Lemma 4.4 and the fact that $\delta_{\tilde{P}}\colon \underline{D}(\tilde{A}) \to \underline{D}(\tilde{A}^*)$ it follows that

$$\delta_{\tilde{P}} = -\int_0^\infty e^{\tilde{A}^* t} c\ell(\tilde{A}^*\delta_{\tilde{P}}+\delta_{\tilde{P}}\tilde{A})e^{\tilde{A}t}\, dt.$$

Now, using the technique of Lemma 4.3 with the role of $\tilde{R}$ played by $-c\ell(\tilde{A}^*\delta_{\tilde{P}}+\delta_{\tilde{P}}\tilde{A})$, we see that

$$\mathrm{tr}[\tilde{Q}c\ell(\tilde{A}^*\delta_{\tilde{P}}+\delta_{\tilde{P}}\tilde{A})] = -\mathrm{tr}[\delta_{\tilde{P}}\tilde{V}] = \mathrm{tr}[\delta_{\tilde{P}}c\ell(\tilde{A}\tilde{Q}+\tilde{Q}\tilde{A}^*)].$$

Similarly, it can be shown that

$$\mathrm{tr}[\tilde{P}c\ell(\tilde{A}\delta_{\tilde{Q}}+\delta_{\tilde{Q}}\tilde{A}^*)] = \mathrm{tr}[\delta_{\tilde{Q}}c\ell(\tilde{A}^*\tilde{P}+\tilde{P}\tilde{A})]. \tag{4.24}$$

Now substitute (4.23) and (4.24) into (4.22) and rearrange the second term in (4.22) so that

$$\begin{aligned}
\Omega &= \tfrac{1}{2}\mathrm{tr}[\delta_{\tilde{Q}}c\ell(\tilde{A}^*\tilde{P}+\tilde{P}\tilde{A}) + \delta_{\tilde{P}}c\ell(\tilde{A}\tilde{Q}+\tilde{Q}\tilde{A}^*)] \\
&\quad + \tfrac{1}{2}\mathrm{tr}[\delta_{\tilde{Q}}(\delta_{\tilde{A}}^*\tilde{P}+\tilde{P}\delta_{\tilde{A}}) + \delta_{\tilde{P}}(\delta_{\tilde{A}}\tilde{Q}+\tilde{Q}\delta_{\tilde{A}}^*)] \\
&\quad - \tfrac{1}{2}\mathrm{tr}[\delta_{\tilde{Q}}c\ell(\tilde{A}'^*\tilde{P}'+\tilde{P}'\tilde{A}') + \delta_{\tilde{P}}c\ell(\tilde{A}'\tilde{Q}'+\tilde{Q}'\tilde{A}'^*)] \\
&= -\tfrac{1}{2}\mathrm{tr}[\delta_{\tilde{Q}}c\ell(\tilde{A}'^*\delta_{\tilde{P}}+\delta_{\tilde{P}}\tilde{A}') + \delta_{\tilde{P}}c\ell(\tilde{A}'\delta_{\tilde{Q}}+\delta_{\tilde{Q}}\tilde{A}'^*)].
\end{aligned}$$

Using (4.8) to obtain

$$0 = \tilde{A}'\delta_{\tilde{Q}} + \delta_{\tilde{Q}}\tilde{A}'^* + \delta_{\tilde{A}}\tilde{Q} + \tilde{Q}\delta_{\tilde{A}}^* + \delta_{\tilde{V}}$$

and (4.10) to obtain a similar relation involving $\tilde{P}$, we have

$$\Omega = \mathrm{tr}[\delta_{\tilde{Q}}(\delta_{\tilde{A}}^{*}\tilde{P}+\tilde{P}\delta_{\tilde{A}}+\delta_{\tilde{R}})] + \mathrm{tr}[\delta_{\tilde{P}}(\delta_{\tilde{A}}\tilde{Q}+\tilde{Q}\delta_{\tilde{A}}^{*}+\delta_{\tilde{V}})].$$

Restricting (A_c', B_c', C_c') to N (see (4.14), using Lemma 4.6 and noting that $\delta_{\tilde{A}}$ and $\delta_{\tilde{R}}$ have finite rank, it follows that there exists $c_1 > 0$ such that

$$||\Omega|| \le c_1 ||(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})||^2. \tag{4.25}$$

Combining Ω with the second-order terms in Λ yields the desired result. □

Lemma 4.8. A_+ is open.

Proof. From the "generic" property of controllability and observability ([62], p. 44) there exists an open neighborhood of (A_c, B_c, C_c) each of whose elements is minimal. Combining this fact with Lemma 4.5 yields the desired result. □

Lemma 4.9. Q_2 and P_2 are positive definite.

Proof. First note that expanding the $\mathbb{R}^{n_c \times n_c}$-component of the Lyapunov equation (4.8) yields (4.50) below. By a minor extension of results from [66] or [67], (4.50) can be rewritten as

$$0 = (A_c+B_cCQ_{12}Q_2^+)Q_2 + Q_2(A_c+B_cCQ_{12}Q_2^+)^T + B_cV_2B_c^T,$$

where Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2.
Next note that since (A_c, B_c) is controllable then so is $(A_c+B_cCQ_{12}Q_2^+, B_cV_2^{\frac{1}{2}})$. Now, since Q_2 and $B_cV_2B_c^T$ are nonnegative definite, it follows from Lemma 12.2 of [62] that Q_2 is positive definite. Similar arguments show that P_2 is positive definite. □

Having established Lemmas 4.1-4.9, we can now proceed with the proof of the Main Theorem. Let $(A_c, B_c, C_c) \in \underline{A}_+$ be as in the Main Theorem and consider (4.19) with (A_c', B_c', C_c') confined to $\underline{A}_+$. Because $\underline{L}: \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \to \mathbb{R}$ is a bounded linear functional and $\underline{A}_+$ is open, the convergence in (4.21) implies that $\underline{L}$ is precisely the Frechet derivative of J with respect to (A_c, B_c, C_c). Since $\underline{A}_+$ is open, the optimality of (A_c, B_c, C_c) implies

$$\underline{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) = 0 \tag{4.26}$$

for all $(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})$. Clearly, (4.26) is equivalent to

$$Z_2 = 0, \tag{4.27}$$

$$V_2 B_c^T P_2 + C Z_{21}^* = 0, \tag{4.28}$$

$$Q_2 C_c^T R_2 + Z_{12}^* B = 0. \tag{4.29}$$

Thus, B_c and C_c are given by

$$B_c = -P_2^{-1} Z_{21} C^* V_2^{-1}, \tag{4.30}$$

$$C_c = -R_2^{-1} B^* Z_{12} Q_2^{-1}. \tag{4.31}$$

Although B_c and C_c are now determined in terms of $\tilde{Q}$ and $\tilde{P}$, A_c remains to be found. Moreover, $\tilde{Q}$ and $\tilde{P}$ themselves depend (via (4.8) and (4.10)) on B_c and C_c. Hence our task now is to consolidate and simplify (4.7)-(4.10), (4.27), (4.30) and (4.31) to obtain the more tractable conditions (3.9)-(3.18). To this end let us define new variables

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^*, \qquad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^*, \tag{4.32a,b}$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^*, \qquad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^*. \tag{4.33a,b}$$

Clearly, $\hat{Q}$ and $\hat{P}$ are nonnegative definite and have finite rank. Since by Lemma 4.2 $\tilde{Q},\tilde{P} \in \underline{B}_1(\underline{\tilde{H}})$, it can be seen that $Q_1, P_1 \in \underline{B}_1(\underline{H})$, which implies $Q,P \in \underline{B}_1(\underline{H})$. To show that Q and P are nonnegative definite, note that Q is the $\underline{B}(\underline{H})$-component of the nonnegative definite, operator $\underline{Q}\tilde{Q}\underline{Q}^* \in \underline{B}(\underline{\tilde{H}})$, where

$$\underline{Q} \triangleq \begin{bmatrix} I_{\underline{H}} & -Q_{12} Q_2^{-1} \\ 0 & -I_{n_c} \end{bmatrix} .$$

Similarly, P is nonnegative definite.

From the domain conditions (4.7) and (4.9) it follows that

$$Q_1\colon \underline{D}(A^*) \to \underline{D}(A), \quad P_1\colon \underline{D}(A) \to \underline{D}(A^*), \tag{4.34a,b}$$

$$Q_{12}\colon \mathbb{R}^{n_c} \to \underline{D}(A), \quad P_{12}\colon \mathbb{R}^{n_c} \to \underline{D}(A^*), \tag{4.35a,b}$$

which lead to (3.12) and (3.13).

Next note that (4.27) is equivalent to (3.8) with

$$G \triangleq Q_2^{-1} Q_{12}^*, \quad \Gamma \triangleq -P_2^{-1} P_{12}^* \tag{4.36a,b}$$

and that (3.7) holds with

$$M \triangleq Q_2 P_2. \tag{4.37}$$

Since Q_2 and P_2 are positive definite, Lemma 2.6 implies that M is positive semisimple. We can also define $\tau = G^*\Gamma$ which, by (3.8) satisfies $\tau^2 = \tau$. It is helpful to note the identities

$$\hat{Q} = Q_{12}G = G^*Q_{12}^*, \quad \hat{P} = -P_{12}\Gamma = -\Gamma^* P_{12}^*, \tag{4.38a,b}$$

$$\hat{Q} = G^*Q_2G, \quad \hat{P} = \Gamma^* P_2\Gamma, \tag{4.39a,b}$$

$$\tau G^* = G^*, \quad \Gamma\tau = \Gamma, \tag{4.40a,b}$$

$$\hat{Q} = \tau\hat{Q}, \quad \hat{P} = \hat{P}\tau, \tag{4.41a,b}$$

$$\hat{Q}\hat{P} = -Q_{12}P_{12}^*. \tag{4.42}$$

From (3.8) and (2.1) it follows that

$$\rho(G) = \rho(\Gamma) = n_c, \tag{4.43a,b}$$

$$\rho(Q_{12}) = \rho(P_{12}) = n_c. \tag{4.44a,b}$$

Hence, (2.2) and (4.38) imply $n_c = \rho(Q_{12}) + \rho(G) - n_c \leq \rho(\hat{Q}) \leq \rho(Q_{12}) = n_c$, which yields (3.14a). Similarly, (3.14b) holds and (3.14c) follows from (2.2) and (4.42).

Using (4.38) and (4.39), the components of $\tilde{Q}$ and $\tilde{P}$ can be written in terms of $G, \Gamma, Q, P, \hat{Q}$ and $\hat{P}$ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \tag{4.45}$$

$$Q_{12} = \hat{Q}\Gamma^*, \quad P_{12} = -\hat{P}G^*, \tag{4.46}$$

$$Q_2 = \Gamma\hat{Q}\Gamma^*, \quad P_2 = G\hat{P}G^*. \tag{4.47}$$

Now (3.10) and (3.11) can be obtained by substituting (4.45)-(4.47) into (4.30) and (4.31).

Expanding the $\underline{B}(\underline{H})$, $\underline{B}(\mathbb{R}^{n_c},\underline{H})$ and $\mathbb{R}^{n_c \times n_c}$ components of (4.8) and (4.10) yields

$$0 = AQ_1 + Q_1A^* + BC_cQ_{12}^* + Q_{12}(BC_c)^* + V_1, \tag{4.48}$$

$$0 = AQ_{12} + Q_{12}A_c^T + BC_cQ_2 + Q_1(B_cC)^*, \tag{4.49}$$

$$0 = A_cQ_2 + Q_2A_c^T + B_cCQ_{12} + Q_{12}^*(B_cC)^* + B_cV_2B_c^T, \tag{4.50}$$

$$0 = A^*P_1 + P_1A + (B_cC)^*P_{12}^* + P_{12}B_cC + R_1, \tag{4.51}$$

$$0 = P_{12}A_c + A^*P_{12} + (B_cC)^*P_2 + P_1BC_c, \tag{4.52}$$

$$0 = A_c^TP_2 + P_2A_c + (BC_c)^*P_{12} + P_{12}^*BC_c + C_c^TR_2C_c. \tag{4.53}$$

Substituting (4.45)-(4.47) into (4.48)-(4.53), using the identities

$$B_cC = \Gamma Q\bar{\Sigma}, \quad BC_c = -\Sigma PG^*,$$

$$B_cV_2B_c^T = \Gamma Q\bar{\Sigma}Q\Gamma^*, \quad C_c^TR_2C_c = GP\Sigma PG^*,$$

and defining

$$A_Q \triangleq A - Q\bar{\Sigma}, \; A_P \triangleq A - \Sigma P,$$

we obtain

$$0 = AQ + QA^* + A_P\hat{Q} + \hat{Q}A_P + V_1, \tag{4.54}$$

$$0 = [A_P\hat{Q} + Q\bar{\Sigma}Q + \hat{Q}(\Gamma^*A_c^TG + \bar{\Sigma}Q)]\Gamma^*, \tag{4.55}$$

$$0 = \Gamma[G^*A_c\Gamma\hat{Q} + Q\bar{\Sigma}\hat{Q} + Q\bar{\Sigma}Q + \hat{Q}(\Gamma^*A_c^TG + \bar{\Sigma}Q)]\Gamma^*, \tag{4.56}$$

$$0 = A^*P + PA + A_Q^*\hat{P} + \hat{P}A_Q + R_1, \tag{4.57}$$

$$0 = -[A_Q^*\hat{P} + P\Sigma P + \hat{P}(G^*A_c\Gamma + \Sigma P)]G^*, \tag{4.58}$$

$$0 = G[\Gamma^*A_c^TG\hat{P} + P\Sigma\hat{P} + P\Sigma P + \hat{P}(G^*A_c\Gamma + \Sigma P)]G^*. \tag{4.59}$$

We are now in a position to determine A_c by computing (4.56) - Γ(4.55) which yields (3.9). Alternatively, A_c can be obtained by computing (4.59) + G(4.58). As mentioned in Section 3, (3.9) is valid since $G^*: \mathbb{R}^{n_c} \rightarrow \underline{D}(A)$ and A_c^T is given by (3.26).

Next we substitute the expressions for A_c and A_c^T into (4.55), (4.56), (4.58) and (4.59) and compute the relations (4.55)G, G^*(4.56)G, -(4.58)Γ and Γ^*(4.59)Γ to obtain, respectively,

$$0 = [A_P\hat{Q} + \hat{Q}A_P^* + Q\bar{\Sigma}Q]\tau^*, \tag{4.60}$$

$$0 = \tau[A_P\hat{Q} + \hat{Q}A_P^* + Q\bar{\Sigma}Q]\tau^*, \tag{4.61}$$

$$0 = [A_Q^*\hat{P} + \hat{P}A_Q + P\Sigma P]\tau, \tag{4.62}$$

$$0 = \tau^*[A_Q^*\hat{P} + \hat{P}A_Q + P\Sigma P]\tau. \tag{4.63}$$

Note that (4.60)-(4.63) are <u>equivalent</u> to (4.55), (4.56), (4.58) and (4.59) since G and Γ have full rank. Since (4.61) = τ(4.60) and (4.63) = τ^*(4.62), (4.61) and (4.63) are superfluous and can be omitted. Thus we have derived (3.17) and (3.18).

To obtain (3.15) and (3.16) we need only compute the relations (4.54) + τ(4.60) - (4.60) - $(4.60)^*$ and (4.57) + τ^*(4.62) - (4.62) - $(4.62)^*$ and use (4.41).

Finally, to show that the preceding development entails no loss of generality in the optimality conditions we now use (3.9)-(3.18) to obtain

(4.7)-(4.10) and (4.27)-(4.29). Let $A_c, B_c, C_c, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}$ be as in the theorem statement and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (4.45)-(4.47). Note that (3.12) and (3.13) imply (4.34) and (4.35) and hence (4.7) and (4.9). Using (3.8), (3.10), (3.11) and (3.22) it is easy to verify (4.27)-(4.29). Finally, substitute (4.32), (4.33) and (4.36) into (3.15)-(3.18), reverse the steps taken earlier in the proof and use (3.9)-(3.11) to obtain (4.8) and (4.10), which completes the proof. □

5. CONCLUDING REMARKS

This paper has considered the problem of quadratically optimal, steady-state, fixed-order dynamic compensation for linear infinite-dimensional systems. The Main Theorem presents the stationarity conditions of the optimization problem in a highly simplified and rigorous form. The "optimal projection equations" (3.15)-(3.18) (or, equivalently, (3.27)-(3.30)) of the Main Theorem reveal the essential structure of the first-order necessary conditions and display the central role played by the optimal projection τ. The relationship of the Main Theorem to the standard finite-dimensional steady-state LQG problem can be demonstrated by replacing τ with the identity matrix and noting that (3.27) and (3.28) reduce immediately to the familiar pair of operator Riccati equations and that (3.29) and (3.30) yield the usual controllability and observability gramians.

Inasmuch as the Main Theorem is a fundamental generalization of classical steady-state LQG theory, a number of issues must be reexamined. Hence, in conclusion we should like to point out some possible extensions of the Main Theorem along with directions for further research.

1. Sufficiency theory. Although sufficient conditions for the existence of an optimal compensator were not investigated in this paper, auxiliary conditions based upon the structure of (3.15)-(3.18) could perhaps be imposed upon Q, P, $\hat{Q}$ and $\hat{P}$ to single out the global optimum from amongst the local minima. This would be similar to the situation in LQG theory where, under stabilizability and detectability hypotheses, optimal stabilizing Q and P are identified as the unique positive semidefinite solutions of the pair of algebraic Riccati equations. Second and higher-order optimality conditions appear promising in this regard and are currently being investigated.

2. Stabilizability. Just as in the full-order LQG problem, one would expect a natural relationship between the structure of the optimal solution and stabilizability/detectability hypotheses. The results of [41], [42] and [68] could serve as a starting point in this regard.

3. Numerical algorithms. In practical situations, the distributed parameter system would be replaced by a high-order discretized model for which the matrix version (rather than the operator version) of the optimal projection equations could be solved numerically. A numerical algorithm for solving the matrix version of the optimal projection equations has been developed in [32] and [34]. The proposed computational scheme is fundamentally quite different from gradient search algorithms ([17,18,21,22,24,25,28,30]) in that it operates through direct solution of the optimal projection equations by iterative refinement of the optimal projection.

4. Convergence. One of the principal uses for the optimal projection equations will be to understand the relationship between fixed-order dynamic-compensator designs which are optimal with respect to approximate models and the optimal fixed-order dynamic compensator for the distributed parameter system itself. By considering a sequence of nth-order approximate models which converge to the distributed parameter system, conditions would be sought guaranteeing that the sequence of fixed-order compensators based on each approximate model approach the optimal dynamic compensator based upon the distributed parameter system (see [38-40]. This approach is analogous to the convergence results obtained in [7,8] with the major difference being that the optimal projection

equations permit the order of the compensator to remain fixed in accordance with real-world implementation constraints whereas in [7-9] the order of the compensator increases without bound.

5. Unbounded control and observation. An important generalization of the problem considered in this paper involves the case in which the input and output operators B and C are unbounded. The mathematical details for this problem are considerably more complex (see, e.g., [69]).

6. Singular observation noise/singular control weighting. As pointed out in [22,33,36] the assumptions of nonsingular control weighting and nonsingular observation noise preclude the use of direct output feedback as in

$$u(t) = C_c x_c(t) + D_c y(t) \tag{5.1}$$

since J is undefined unless

$$\operatorname{tr}[D_c^T R_2 D_c V_2] = 0 \quad (\Leftrightarrow \; R_2 D_c V_2 = 0).$$

Although with due attention to (5.1) direct output feedback can be used in the singular case, the nature of the problem forebodes all of the difficulties associated with the singular LQG problem. Note that the deterministic output feedback problem ([70]), when viewed in this context, is highly singular.

7. Discrete-time system/discrete-time compensator. Digital implementation can be modelled by a discrete-time compensator with control of a continuous-time system facilitated by sampling and reconstruction devices. See [71] for results in this direction.

8. Cross weighting/correlated disturbance and observation noise. This extension is straightforward and entirely analogous to the LQG case (see, e.g., [18], p. 351).

ACKNOWLEDGMENT

We wish to thank Ardeth P. Grant for excellent and careful typing of the original manuscript version of this paper. Word processing support for the revised paper was provided by Harris Corporation, GASD.

REFERENCES

1. M. J. Balas, Toward a more practical control theory for distributed parameter systems, in Control and Dynamic Systems: Advances in Theory and Applications, 19(1982), C. T. Leondes, ed., Academic Press, New York.
2. M. Athans, Toward a practical theory of distributed parameter systems, IEEE Trans. Autom. Contr., AC-15(1970), pp. 245-247.
3. S. A. Reible, Acoustoelectric convolver technology for spread-spectrum communications, IEEE Trans. Microwave Theory Tech., MTT-29(1981), pp. 463-473.
4. R. F. Curtain and A. J. Pritchard, Infinite Dimensional Linear Systems Theory, Springer-Verlag, New York, 1978.
5. J. S. Gibson, The Riccati integral equations for optimal control problems on Hilbert spaces, SIAM J. Contr. Optim., 17(1979), pp. 537-565.
6. A. V. Balakrishnan, Applied Functional Analysis, Springer-Verlag, New York, 1981.
7. J. S. Gibson, An analysis of optimal modal regulation: conve -ence and stability, SIAM J. Contr. Optim., 19(1981), pp. 686-707.
8. J. S. Gibson, Linear-quadratic optimal control of hereditary differential systems: infinite dimensional Riccati equations and numerical approximations, SIAM J. Contr. Optim., 21(1983), pp. 95-139.

8a. H. T. Banks and K. Kunisch, The linear regulator problem for parabolic systems, SIAM J. Contr. Optim. 22(1984), pp. 684-698.

9. H. T. Banks, K. Ito and I. G. Rosen, A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations, ICASE Report 82-31, Institute for Computer Applications in Science and Engineering, Hampton, VA, 1982; SIAM J. Sci. Stat. Comput., 5(1984), to appear.

10. M. Aoki, Control of large-scale dynamic systems by aggregation, IEEE Trans. Autom. Contr., AC-13(1968), pp. 246-253.

11. D. A. Wilson, Optimum solution of model-reduction problem, Proc. IEE, 117(1970), pp. 1161-1165.

12. B. C. Moore, Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Trans. Autom. Contr., AC-26(1981), pp. 17-32.

13. R. E. Skelton, Component cost analysis of large scale systems, in Control and Dynamic Systems, 18(1981), edited by C. T. Leondes, Academic Press.

14. E. A. Jonckheere and L. M. Silverman, A new set of invariants for linear systems - application to reduced-order compensator design, IEEE Trans. Autom. Contr., AC-28(1983), pp. 953-964.

15. A. Yousuff and R. E. Skelton, Controller reduction by component cost analysis, IEEE Trans. Autom. Contr., AC-29(1984), pp. 520-530.

16. T. L. Johnson and M. Athans, On the design of optimal constrained dynamic compensators for linear constant systems, IEEE Trans. Autom. Contr., AC-15(1970), pp. 658-660.

17. W. S. Levine, T. L. Johnson and M. Athans, Optimal limited state variable feedback controllers for linear systems, IEEE Trans Autom. Contr., AC-16(1971), pp. 785-793.

18. K. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, New York, 1972.

19. D. B. Rom and P. E. Sarachik, The design of optimal compensators for linear constant systems with inaccessible states, IEEE Trans. Autom. Contr., AC-18(1973), pp. 509-512.

20. M. Sidar and B.-Z. Kurtaran, Optimal low-order controllers for linear stochastic systems, Int. J. Contr., 22(1975), pp. 377-387.

21. J. M. Mendel and J. Feather, On the design of optimal time-invariant compensators for linear stochastic time-invariant systems, IEEE Trans. Autom. Contr., AC-20(1975), pp. 653-657.

22. S. Basuthakur and C. H. Knapp, Optimal constant controllers for stochastic linear systems, IEEE Trans. Autom. Contr., AC-20(1975), pp. 664-666.

23. R. B. Asher and J. C. Durrett, Linear discrete stochastic control with a reduced-order dynamic compensator, IEEE Trans. Autom. Contr., AC-21(1976), pp. 626-627.

24. W. J. Naeije and O. H. Bosgra, The design of dynamic compensators for linear multivariable systems, 1977 IFAC, Fredricton, NB, Canada, pp. 205-212.

25. H. R. Sirisena and S. S. Choi, Design of optimal constrained dynamic compensators for non-stationary linear stochastic systems, Int. J. Contr., 25(1977), pp. 513-524.

26. P. J. Blanvillain and T. L. Johnson, Specific-optimal control with a dual minimal-order observer-based compensator, Int. J. Contr., 28(1978), pp. 277-294.

27. P. J. Blanvillain and T. L. Johnson, Invariants of optimal minimal-order observer based compensators, IEEE Trans. Autom. Contr., AC-23(1978), pp. 473-474.

28. C. J. Wenk and C. H. Knapp, Parameter optimization in linear systems with arbitrarily constrained controller structure, IEEE Trans. Autom. Contr., AC-25(1980), pp. 496-500.

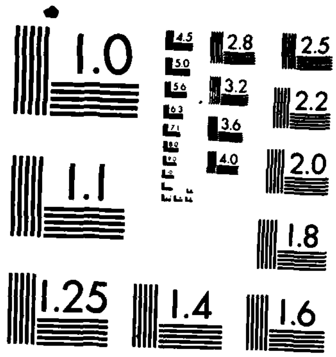
AD-A161 355 EXPLORATION OF THE MAXIMUM ENTROPY/OPTIMAL PROJECTION APPROACH TO CONTROL (U) HARRIS CORP MELBOURNE FL GOVERNMENT AEROSPACE SYSTEMS DIV D C HYLAND ET AL 3/3

UNCLASSIFIED FEB 85 AFOSR-TR-85-0966 F49620-84-C-0015 F/G 13/13 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

29. J. O'Reilly, Optimal low-order feedback controllers for linear discrete-time systems, in Control and Dynamic Systems, 16(1980), edited by C. T. Leondes, Academic Press.

30. D. P. Looze and N. R. Sandell, Jr., Gradient calculations for linear quadratic fixed control structure problems,' IEEE Trans. Autom. Contr., AC-25(1980), pp. 285-288.

31. D. C. Hyland, Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters, AIAA 20th Aerospace Sciences Mtg., Orlando, FL, Jan. 1982.

32. D. C. Hyland, The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results, AIAA 21st Aerospace Sciences Mtg., Reno, NV, Jan. 1983.

33. D. C. Hyland and D. S. Bernstein, Explicit optimality conditions for fixed-order dynamic compensation, IEEE Conf. Dec. Contr., San Antonio, TX, Dec. 1983.

34. D. C. Hyland, Comparison of various controller-reduction methods: Suboptimal versus optimal projection, AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

35. D. S. Bernstein and D. C. Hyland, The optimal projection equations for fixed-order dynamic compensation of distributed parameter systems, AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

36. D. C. Hyland and D. S. Bernstein, The optimal projection equations for fixed-order dynamic compensation, IEEE Trans. Autom. Contr. (to appear).

37. D. C. Hyland and D. S. Bernstein, The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore, IEEE Conf. Dec. Contr. Las Vegas, NV, Dec 1984.

38. T. L. Johnson, Optimization of low order compensators for infinite dimensional systems, Proc. of 9th IFIP Symp. on Optimization Techniques, Warsaw, Poland, September 1979.

39. R. K. Pearson, Optimal fixed-form compensators for large space structures, in ACOSS SIX (Active Control of Space Structures), RADC-TR-81-289, Final Technical Report, Rome Air Development Center, Griffiss AFB, New York, 1981.

40. R. K. Pearson, Optimal velocity feedback control of flexible structures, Ph.D. Dis., M.I.T. Dept. Elec. Eng. Comp. Sci., 1982.

41. R. F. Curtain, Compensators for infinite-dimensional linear systems, J. Franklin Inst., 315(1983), pp. 331-346.

42. J. M. Schumacher, A direct approach to compensator design for distributed parameter systems, SIAM J. Contr. Optim., 21(1983), pp. 823-836.

43. D. L. Russel, Linear stabilizaton of the linear oscillator in Hilbert space, J. Math. Anal. Appl., 25(1969), pp. 663-675.

44. D. L. Russel, Decay rates for weakly damped systems in Hilbert space obtained with control theoretic methods, J. Diff. Eqns., 19(1975), pp. 344-370.

45. M. J. Balas, Modal control of certain flexible dynamic systems, SIAM J. Contr. Optim., 16(1978), pp. 450-462.

46. M. J. Balas, Feedback control of flexible systems, IEEE Trans. Autom. Contr., AC-23(1978), pp. 673-679.

47. J. S. Gibson, A note on stabilization of infinite dimensional linear oscillators by compact feedback, SIAM J. Contr. Optim., 18(1980), pp. 311-316.

48. M. J. Balas, Trends in large space structure control theory: Fondest hopes, wildest dreams, IEEE Trans. Autom. Contr., AC-24(1982), pp. 522-535.

49. T. L. Johnson, Progress in modelling and control of flexible spacecraft, J. Franklin Inst., 315(1983), pp. 495-520.

49a. M. J. Balas, Feedback control of dissipative hyperbolic distributed parameter systems with finite dimensional controllers, J. Math. Anal. Appl., 98(1984), pp. 1-24.

50. T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.

51. J. R. Ringrose, Compact Non-Self-Adjoint Operators, Van Nostrand Reinhold Co., London, 1971.

52. I. C. Gohberg and M. G. Krein, Introduction to the theory of linear nonselfadjoint operators, Translations of Mathematical Monographs, Vol. 18, American Mathematical Society, Providence, RI, 1966.

53. I. Gohberg and S. Goldberg, Basic Operator Theory, Birkhauser, Boston, 1981.

54. F. R. Gantmacher, The Theory of Matrices, Vol. I, Chelsea, NY, 1977.

55. C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Its Applications, John Wiley and Sons, New York, 1971.

56. B. Noble and J. W. Daniel, Applied Linear Algebra, Second Edition, Prentice-Hall, Englewood Cliffs, NJ, 1977.

57. R. F. Curtain and A. J. Pritchard, Functional Analysis in Modern Applied Mathematics, Academic Press, London, 1977.

58. S. Chakrabarti, B. B. Battacharyya and M. N. S. Swamy, On simultaneous diagonalization of a collection of hermitian matrices, The Matrix and Tensor Quarterly, 29(1978), pp. 35-54.

59. C. T. Mullis and R. A. Roberts, Synthesis of minimum roundoff noise fixed point digital filters, IEEE Trans. Circ. Syst., CAS-23(1976), pp. 551-562.

60. A. J. Laub, Computation of balancing transformation, Proc. 1980 Joint Autom. Contr. Conf., San Francisco, CA, Aug. 1980.

61. E. Jonckheere, Open-loop and closed loop approximations of linear systems and associated balanced realizations, 1982 Symp. Circ. Sys., Rome, Italy, May 1982.

62. W. M. Wonham, Linear Multivariable Control: A Geometric Approach, Springer-Verlag, 1974.

63. D. C. Lay, Spectral properties of generalized inverses of linear operators, SIAM J. Appl. Math., 29(1975), pp. 103-109.

64. S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

65. P. Robert, On the group-inverse of a linear transformation, J. Math. Anal. Appl., 22(1968), pp. 658-669.

66. A. Albert, Conditions for positive and nonnegative definiteness in terms of pseudo inverse, SIAM J. Appl. Math., 17(1969), pp. 434-440.

67. E. Kreindler and A. Jameson, Conditions for nonnegativeness of partitioned matrices, IEEE Trans. Autom. Contr., AC-17(1972), pp. 147-148.

68. C. N. Nett, C. A. Jacobson and M. J. Balas, Fractional representation theory: Robustness results with applications to finite dimensional control of a class of linear distributed systems, IEEE Conf. Dec. Contr., San Antonio, TX, Dec 1983.

69. R. F. Curtain, Finite-dimensional compensators for parabolic distributed systems with unbounded control and observation, SIAM J. Contr. Optim., 22(1984), pp. 255-276.

70. W. S. Levine and M. Athans, On the determination of the optimal constant output feedback gains for linear multivariable systems, IEEE Trans. Autom. Contr., AC-15(1970), pp. 44-48.

71. M. J. Balas, The structure of discrete-time finite-dimensional control of distributed parameter systems, Proc. IEEE Int. Large Scale Systems Symp., Virginia Beach, VA, 1982.

APPENDIX E

(REFERENCE [31])

IFAC WORKSHOP

ON

MODEL ERROR CONCEPTS AND COMPENSATION

JUNE 17, 18, 1985

BOSTON, MASSACHUSETTS

OPTIMAL PROJECTION/MAXIMUM ENTROPY STOCHASTIC MODELLING AND REDUCED-ORDER DESIGN SYNTHESIS

(ABSTRACT)

by

DENNIS S. BERNSTEIN and DAVID C. HYLAND

Controls Analysis and Synthesis Group
Government Aerospace Systems Division
Harris Corporation
Melbourne, Florida 32902

1. Introduction

This paper will be broad in scope and has the following objectives:

1. to discuss the underlying philosophy and motivation of the optimal projection/maximum entropy (OP/ME) stochastic modelling and reduced-order design methodology for high-order systems with parameter uncertainties;

2. to present a rigorous mathematical development of the principal design results for reduced-order modelling, reduced-order state estimation and reduced-order dynamic compensation including the effects of parameter uncertainties; and

3. to contrast this approach philosophically and technically with several alternative methods with regard to capabilities and limitations.

The basis for this paper is references [1-25] along with recently-obtained results.

The OP/ME approach, as its name *suggests*, represents the *synthesis of* two distinct ideas: (1) reduced-order system design, e.g., model, state estimator or dynamic compensator, for a given high-order plant (i.e., optimal projection design) and (2) minimum-information stochastic modelling of parameter uncertainties (i.e., maximum entropy modelling.) In view of the theme of the workshop, the overwhelming emphasis of this paper will be on maximum entropy modelling. However, since order uncertainties due to inadvertent or intentional system truncation (either from infinite-to-finite or finite-to-finite dimensions) at various stages in the design process are also pertinent to the workshop, the reduced-order aspect of our work will also be included. Maximum entropy modelling is discussed in [1-13,15] and optimal projection design is studied in [6,10,12,14,16-25]. Before *discussing* maximum entropy modelling, we shall briefly review optimal projection design for high-order systems, assuming for the moment an absence of modelling errors. Although mathematically classical, optimal projection design represents a novel approach to the reduced-order modelling, reduced-order state-estimation and reduced-order dynamic-compensation problems.

This approach has recently been vigorously developed and it should be noted that [22] is scheduled for archival publication and [23-25] are currently under similar review.

The optimal projection approach is based entirely on a series of three theorems (these can be obtained as corollaries of results given below) which characterize the quadratically optimal reduced-order model, state estimator and dynamic compensator. Assuming a purely dynamic linear system structure for the desired system (model, estimator or compensator) whose order is determined by implementation constraints (e.g., reliability, complexity or computing capability), a parameter optimization approach is taken. There is, of course, nothing novel about this approach per se and it has been widely studied in the control literature [26-39]. Although the model-reduction problem was also handled in this way in [40-42], to the authors' knowledge the widely-studied reduced-order state-estimation problem ([43,44]) was not directly approached in this fashion. Clearly, the parameter optimization approach fell into disrepute because of the extreme complexity of the grossly unwieldy first-order necessary conditions which afforded little insight and engendered brute force gradient search techniques. The crucial discovery occurred in [6] where it was revealed that the necessary condition for the dynamic-compensation problem give rise to the definition of an <u>optimal projection</u> as a rigorous, unassailable consequence of quadratic optimality without recourse to <u>ad hoc</u> methods as in [45-54]. Exploitation of this projection leads to immense simplification of the "primitive" form of the necessary conditions for each of these problems. As summarized in Figure 1, the modelling, estimation and compensation design equations form a natural progression: a coupled system of 2, 3 or 4 matrix equations whose solutions determine the desired gains (A_m, B_m, C_m), (A_e, B_e, C_e) or (A_c, B_c, C_c). The novel equations are the modified Lyapunov equations for the reduced-order modelling problem, versions of which arise in the estimation and compensation problems. Since the modified Riccati equations are analogous to the standard observer and regulator Riccati equations, the optimal projection equations for the reduced-order state-estimation and dynamic-compensation problems appear to provide a fundamental generalization of steady-state Kalman filter and LQG theory.

Although optimal projection design deals directly and rigorously with the question of system dimension by trading order off against performance, it is, nevertheless, predicated upon the availability of a completely accurate plant and disturbance model. Maximum entropy modelling, however, addresses the robustness problem by permitting direct inclusion of parameter uncertainties in the plant and disturbance models so that optimal projection design plus maximum entropy modelling automatically yields system designs (reduced-order models, state estimators and dynamic compensators) that trade performance off against modelling uncertainties. In maximum entropy modelling, uncertainties are modelled at their a priori levels and there is no adaptation or learning provided for in the design, i.e., the control is nondual ([55-57]).

Before attempting an overview of the maximum entropy approach it is important to discuss the class of problems that motivated this work, namely, control of large flexible space structures. A finite-element model of a large flexible space structure is, generally, an extremely high-order system. For example, a version of the widely-studied Draper Model #2 includes 150 modes and 6 disturbance states, i.e., a total of 306 states, along with 9 sensors and 9 actuators. The size of the model and the coupling between sensors and actuators renders classical control-design methods useless and all but confounds attempts to use LQG to obtain a controller of manageable order. Indeed, these difficulties were a prime motivation for the optimal projection approach. Besides the high order of these systems, finite element modelling is known to have poor accuracy, particularly for the high-order modes. Reasonable and not overly-conservative uncertainty estimates predict 30-50 percent error in modal frequencies after the first ten modes, with the situation considerably more complex (and pessimistic) for damping estimates. Otherwise-successful control-design methodologies widely promulgated in the aerospace community were severely strained in the face of such difficulties.

Maximum entropy modelling is a form of stochastic modelling. Although external disturbances are traditionally modelled stochastically as random processes, the use of stochastic theory to model plant parameter uncertainty has seen relatively limited application. We seek to dispel all objections to a stochastic parameter uncertainty model by invoking the modern information-theoretic interpretation of probability theory. Rather than regarding the probability of an event as the limiting frequency of numerous repetitions (as, e.g., the number of heads in 1,000 coin tosses) we adopt the view that the

probability of an event is a subjective quantity which reflects the observer's certainty as to whether a particular event will or will not occur. This quantity is nothing more than a measure of the information (including, e.g., all theoretical analysis and empirical data) available to the observer. In this sense the validity of a stochastic model of a flexible space structure, for example, does not rely upon the existence of a fleet of such objects (substitute "ensemble" for "fleet" in the classical terminology) but rather resides in the interpretation that it expresses the engineer's certainty or uncertainty regarding the values of physical parameters such as stiffnesses of structural components. This view of probability theory has its roots in Shannon's information theory but was first articulated unambiguously by Jaynes (see [58-61]).*

The preeminent problem in modelling the real world is thus the following: given limited (incomplete) *a priori* data, how can a well-defined (complete) probability model be constructed which is consistent with the available data but which *avoids inventing data which does not exist*? To this end we invoke Jaynes' Maximum Entropy Principle: First, define a measure of ignorance in terms of the information-theoretic entropy, and then determine the probability distribution which *maximizes* this measure subject to agreement with the available data.** The reasoning behind this principle is that the probability distribution which maximizes *a priori* ignorance *must* be the least presumptive (i.e., least likely to invent data) on the average since the amount of *a posteriori* learned information (should all uncertainty suddenly disappear) would necessarily be maximized. If, for some probability distribution, the *a priori* ignorance and hence the *a posteriori* learning were *less* than their maximum value then this distribution must be based upon invented and hence, generally incorrect, data. The Maximum Entropy Principle is clearly desirable for control-system design where the introduction of false data is to be assiduously avoided.

It is shown in [1] that the stochastic model induced by the Maximum Entropy Principle of Jaynes is a Stratonovich multiplicative white noise model. Rather than base this paper on a demonstration of this result, whose derivation

---

*The reader is encouraged to peruse [60] (which is also available in [61]) where this interpretation of probability theory is discussed.

**The smallest collection of data for which a well-defined probability model (called the minimum information model) can be constructed is known as the minimum data set.

is confined to modal (structural) systems, we shall arbitrarily adopt the Stratonovich multiplicative white noise model and proceed by exploring its consequences for general systems. The idea of inducing a stochastic model from limited data remains, however, fundamental to maximum entropy modelling and figures prominently in the results below.

A review of the mathematical and control systems literature on multiplicative white noise is absolutely essential for communicating the contribution of maximum entropy modelling. The theory of stochastic differential equations was placed on a firm mathematical foundation by Ito ([62]) and has been widely developed and applied to modelling, estimation and control problems ([63-91]). The basic linear multiplicative white noise model is given by

$$\dot{\tilde{x}}(t) = (\tilde{A} + \sum_{i=1}^{p} \alpha_i(t)\tilde{A}_i)\tilde{x}(t) + \tilde{w}(t), \tag{1.1}$$

where $\tilde{x}(t) \in \underline{\underline{R}}^{\tilde{n}}$, $\tilde{A}$, $\tilde{A}_i \in \underline{\underline{R}}^{\tilde{n}\times\tilde{n}}$, $\tilde{w}(t)$ is zero-mean Gaussian white disturbance noise with nonnegative-definite intensity $\tilde{V}$, and $\alpha_i(t)$ are zero-mean, unit-intensity Gaussian white noise processes which are mutually uncorrelated and uncorrelated with $\tilde{w}(t)$. The multiplicative white noise model (1.1) can be regarded as a parameter uncertainty model where each $\alpha_i(t)$ corresponds to a single uncertain parameter whose <u>pattern</u> and <u>magnitude</u> are given by $\tilde{A}_i/||\tilde{A}_i||$ and $||\tilde{A}_i||$, respectively.

To see why (1.1) is a minimum information model of parameter uncertainty, note that when the pattern $\tilde{A}_i/||\tilde{A}_i||$ of an uncertain parameter is known, all available data (theoretical and empirical) can be brought to bear ("boiled down") to determine its magnitude $||\tilde{A}_i||$. Clearly, the collection of magnitudes constitutes the minimum data set needed to render (1.1) well defined. For the harmonic oscillator with uncertain natural frequency, the uncertainty magnitude is given by the reciprocal of the <u>relaxation time</u> (see Figure 2).

To eliminate the white noise formalism, the model (1.1) is usually rigorized by the Ito differential equation

$$d\tilde{x}_t = (\tilde{A}dt + \sum_{i=1}^{p} d\alpha_{it}\tilde{A}_i)\tilde{x}_t + d\tilde{w}_t, \tag{1.2}$$

where $d\alpha_{it}$ and dw_t are Brownian motions, i.e., Wiener processes. Although such models were studied extensively for estimator and control design ([72-88]), this approach fell into disrepute with the publication of [90,91] where it was shown for discrete-time systems that sufficiently high uncertainty levels (i.e., magnitudes $||\tilde{A}_i||$ above a threshold) lead to the nonexistence of a steady state solution. Although it was purported in [90] that this "phenomenon" was an "obvious" consequence of high uncertainty levels, these conclusions failed to take into account (possibly because of the discrete-time setting) the subtle relationship between the ordinary differential equation (1.1) and the stochastic differential equation (1.2). Indeed, it was shown in [63] that if a stochastic differential equation is regarded as the limit of a sequence of ordinary differential equations, then (1.2) is <u>not</u> the correct version of (1.1). Instead, the ordinary differential equation (1.1) with multiplicative white noise corresponds to the <u>corrected</u> Ito differential equation

$$dx_t = (\tilde{A}_s dt + \sum_{i=1}^{p} d\alpha_{it}\tilde{A}_i)\tilde{x}_t + d\tilde{w}_t, \tag{1.3}$$

where

$$\tilde{A}_s \triangleq \tilde{A} + \frac{1}{2}\sum_{i=1}^{p} \tilde{A}_i^2 \tag{1.4}$$

which differs from the "naive" equation (1.2) by a systematic drift term. Although skepticism regarding this unusual result was admitted to in [63], the form of (1.3) was corroborated completely independently by Stratonovich in [64], whose results actually appeared in the Russian literature prior to 1965. His approach is based upon an alternative definition of stochastic integration which differs from Ito stochastic integration by a mathematical technicality. The Stratonovich approach, it should be noted, has the interesting feature that approximating sums involve future values of a Brownian motion process which, although physically unacceptable in the classical view of probability, is completely consistent with the information-theoretic interpretation.

In spite of the glaring technicality of the <u>Stratonovich correction</u>, almost all research on the estimation and control of such systems failed to perceive its physical significance. To the authors' knowledge, the work of Gustafson and Speyer [88] was the only paper prior to the appearance of [1] which demonstrated the crucial feature: The Stratonovich correction neutralizes the threshold uncertainty principle! We shall now proceed to demonstrate this fact by means of a compelling example.

First, suppose that zero-point deviations of $\widetilde{x}(t)$ are of interest and are evaluated according to

$$J = \lim_{t\to\infty} \widetilde{x}(t)^T\widetilde{R}\widetilde{x}(t) = \lim_{t\to\infty} \operatorname{tr} \widetilde{Q}(t)\widetilde{R}, \tag{1.5}$$

where $\widetilde{R} \in \underline{\underline{R}}^{\widetilde{n}\times\widetilde{n}}$ and

$$\widetilde{Q}(t) \triangleq \underline{\underline{E}}[\widetilde{x}(t)\widetilde{x}(t)^T]. \tag{1.6}$$

As will be seen, (1.5) can arise in either reduced-order modelling, reduced-order state estimation or reduced-order dynamic-compensation problems. The obvious fact cannot be over emphasized that the sole state statistic of design interest is the state covariance (1.6). From Ito calculus it follows that $\widetilde{Q}(t)$ is given for the naive model (1.2) by

$$\dot{\widetilde{Q}}(t) = \widetilde{A}\widetilde{Q}(t) + \widetilde{Q}(t)\widetilde{A}^T + \sum_{i=1}^{p} \widetilde{A}_i\widetilde{Q}(t)\widetilde{A}_i + \widetilde{V} \tag{1.7}$$

and for the corrected model (1.3) by

$$\dot{\widetilde{Q}}(t) = \widetilde{A}_s\widetilde{Q}(t) + \widetilde{Q}(t)\widetilde{A}_s^T + \sum_{i=1}^{p} \widetilde{A}_i\widetilde{Q}(t)\widetilde{A}_i + \widetilde{V}. \tag{1.8}$$

Each of these stochastic Lyapunov differential equations should be regarded as $\widetilde{n}(\widetilde{n}+1)/2$ ordinary differential equations. The question we wish to address is the following: How do the solutions of the stochastic Lyapunov equations (1.7) and (1.8) differ from each other and the "deterministic" Lyapunov equation

$$\dot{\widetilde{Q}}(t) = \widetilde{A}\widetilde{Q}(t) + \widetilde{Q}(t)\widetilde{A}^T + \widetilde{V}, \tag{1.9}$$

particularly in the presence of high uncertainty levels? The answer to this question of course depends upon the stochastic modification terms which for the naive model are given by

$$\underline{M}_N[\widetilde{Q}(t)] \triangleq \sum_{i=1}^{p} \widetilde{A}_i\widetilde{Q}(t)\widetilde{A}_i^T \tag{1.10}$$

and for the corrected model by

$$\underline{M}_C[\widetilde{Q}(t)] = \sum_{i=1}^{p} [\tfrac{1}{2}\widetilde{A}_i^2\widetilde{Q}(t) + \tfrac{1}{2}\widetilde{Q}(t)\widetilde{A}_i^{2T} + \widetilde{A}_i\widetilde{Q}(t)\widetilde{A}_i^T]. \qquad (1.11)$$

We now consider a system consisting of a pair of lightly-damped modes so that

$$\widetilde{A} = \begin{bmatrix} 0 & -\omega_1 & 0 & 0 \\ \omega_1 & -2\eta_1 & 0 & 0 \\ 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & \omega_2 & -2\eta_2 \end{bmatrix},$$

where $\eta_i \triangleq \zeta_i\omega_i$. To represent frequency uncertainties let

$$\widetilde{A}_1 = \gamma_1 \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \widetilde{A}_2 = \gamma_2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

where for simplicity we have ignored the effects of frequency uncertainties on the effective decay rate η_i. The magnitudes of the uncertainties are scaled by means of γ_1 and γ_2. For this example the naive stochastic modification has the form

$$\underline{M}_N[\widetilde{Q}(t)] = \begin{bmatrix} \gamma_1^2\widetilde{Q}_{22}(t) & -\gamma_1^2\widetilde{Q}_{12}(t) & 0 & 0 \\ -\gamma_1^2\widetilde{Q}_{12}(t) & \gamma_1^2\widetilde{Q}_{11}(t) & 0 & 0 \\ 0 & 0 & \gamma_2^2\widetilde{Q}_{44}(t) & -\gamma_2^2\widetilde{Q}_{34}(t) \\ 0 & 0 & -\gamma_2^2\widetilde{Q}_{34}(t) & \gamma_2^2\widetilde{Q}_{33}(t) \end{bmatrix}.$$

Although the off-diagonal terms have a stabilizing effect, it is clear that the diagonal elements destabilize the state variances. Hence, it is not surprising that for sufficiently high uncertainty levels, i.e., $\gamma_i \gg 0$, the naive model is second-moment unstable. These observations are completely in accordance with the threshold uncertainty principle. The corrected stochastic modification, however, has the form

$$\underline{M}_C[\tilde{Q}(t)] = \begin{bmatrix} \gamma_1^2[\tilde{Q}_{22}(t)-\tilde{Q}_{11}(t)] & -2\gamma_1^2\tilde{Q}_{12}(t) & -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{13}(t) & -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{14}(t) \\ -2\gamma_1^2\tilde{Q}_{12}(t) & \gamma_1^2[\tilde{Q}_{11}(t)-\tilde{Q}_{22}(t)] & -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{23}(t) & -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{24}(t) \\ -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{31}(t) & -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{32}(t) & \gamma_2^2[\tilde{Q}_{44}(t)-\tilde{Q}_{33}(t)] & -2\gamma_2^2\tilde{Q}_{34}(t) \\ -\frac{1}{2}(\gamma_1^2+\gamma_2^2)\tilde{Q}_{41}(t) & -\frac{1}{2}(\gamma_1^2+\gamma_1^2)\tilde{Q}_{42}(t) & -2\gamma_2^2\tilde{Q}_{34}(t) & \gamma_2^2[\tilde{Q}_{33}(t)-\tilde{Q}_{44}(t)] \end{bmatrix}$$

which also has stabilizing off-diagonal elements but has fundamentally different diagonal elements: Rather than destabilizing the state variances, the diagonal elements of the corrected stochastic modification are equilibrating. This effect is even more *striking when* $\underline{M}_N$ and $\underline{M}_C$ are transformed into the basis with respect to which

$$\tilde{A} = \begin{bmatrix} -j\omega_1-\eta_1 & 0 & 0 & 0 \\ 0 & j\omega_1-\eta_1 & 0 & 0 \\ 0 & 0 & -j\omega_2-\eta_2 & 0 \\ 0 & 0 & 0 & j\omega_2-\eta_2 \end{bmatrix},$$

where higher-order terms in η have been ignored. In this basis, the diagonal terms of $\underline{M}_N[\tilde{Q}(t)]$ are destabilizing whereas the diagonal terms of $\underline{M}_C[\tilde{Q}(t)]$ exactly vanish.

The negative coefficients in the off-diagonal terms imply progressive decorrelation between pairs of dynamical states. This informational or statistical damping phenomenon is a direct result of parameter uncertainties that

is captured by the multiplicative white noise model.* The Stratonovich correction, moreover, is crucial: By neutralizing the threshold uncertainty principle, it permits the consideration of long-term effects for arbitrary uncertainty levels.

The far-reaching ramifications of these observations are explored extensively in [1-10]. As an example, assume (as is usually the case in practice) that uncertainties in modal frequency obtained from a finite-element analysis of a large flexible space structure increase with mode number. From the form of $M_C[\tilde{Q}(t)]$ it is easy to deduce that the steady-state covariance

$$\tilde{Q} \triangleq \lim_{t\to\infty} \tilde{Q}(t)$$

satisfying

$$0 = \tilde{A}_s\tilde{Q} + \tilde{Q}\tilde{A}_s^T + \sum_{i=1}^{p} \tilde{A}_i\tilde{Q}\tilde{A}_i^T + \tilde{V} \tag{1.12}$$

becomes increasingly diagonally dominant with increasing frequency and thus assumes the qualitative form given in Figure 3. The benefits of this sparse form are important: The computational effort required to determine the steady-state covariance (and thus to design a closed-loop controller, for example) is directly proportional to the amount of information reposed in the model or, equivalently, inversely proportional to the level of modelled parameter uncertainty. This casts new light on the computational design burden vis-a-vis the modelling question: The computational burden depends only upon the information actually available. A simple control-design exercise involving full-state feedback illustrates this point. The gains for the higher-order modes of the beam in Figure 4, whose frequency uncertainties increase linearly with frequency, were obtained with modest computational effort in spite of $\tilde{n} = 100$ (see Figure 5). Another important ramification of the qualitative form of $\tilde{Q}$ is the automatic generation of a high/low-authority control law. Note that for the higher-order and hence highly-uncertain modes the control gains indicate an inherently-stable, low performance rate-feedback control law, whereas for the lowest-order modes the control law is high authority, i.e., "LQ" in character.

---

*Needless to say, these qualitative effects are statistical and do not refer to what a particular structure sitting in a test facility or on-orbit is actually doing!

2. Design Results

The major design results are the first-order necessary conditions for optimality in the presence of maximum entropy modelling. For the reduced-order modelling problem, the following notation is needed: $x \in \underline{\underline{R}}^n$, $y \in \underline{\underline{R}}^\ell$; $w \in \underline{\underline{R}}^m$ is white noise with intensity $V > 0$; $n_m \leq n$, $x_m \in \underline{\underline{R}}^{n_m}$, $y_m \in R^\ell$, $A, A_1, \ldots, A_p, B, C, A_m, B_m, C_m$ and R where $R > 0$ are real matrices whose dimensions are consistent with their context below; and $\alpha_1, \ldots, \alpha_p$ are zero-mean unit-variance mutually uncorrelated scalar white noises which are also uncorrelated with w. In this problem, the order n_m and structure of the reduced-order model are fixed and the problem is concerned with determining A_m, B_m and C_m.

Optimal Reduced-Order Modelling Problem (ROM). Given the model

$$\dot{x} = (A + \sum_{i=1}^{p} \alpha_i A_i)x + Bw, \tag{2.1}$$

$$y = Cx \tag{2.2}$$

design a reduced-order model

$$\dot{x} = A_m x_m + B_m w, \tag{2.3}$$

$$y_m = C_m x_m \tag{2.4}$$

which minimizes the model-reduction criterion

$$J(A_m, B_m, C_m) = \lim_{t \to \infty} \underline{\underline{E}}[(y-y_m)^T R(y-y_m)]. \tag{2.5}$$

For the reduced-order state-estimation problem, let x, y, A, C and $\alpha_1, \ldots, \alpha_p$ be as in the reduced-order modelling problem. Furthermore, assume $C_1, \ldots, C_p$ have the same dimensions as C; $w_1 \in \underline{\underline{R}}^n$ and $w_2 \in \underline{\underline{R}}^\ell$ are white noises uncorrelated with the α_i's with intensities $V_1 \geq 0$ and $V_2 > 0$, respectively, and cross intensity $V_{12} \in \underline{\underline{R}}^{n \times \ell}$; $n_e \leq n$; $x_e \in \underline{\underline{R}}^{n_e}$, $y_e \in \underline{\underline{R}}^q$, $R \in \underline{\underline{R}}^{q \times q}$; and A_e, B_e and C_e are matrices of suitable dimension.

In this formulation the matrix L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is considered fixed and is determined by implementation constraints, i.e., by the computing capability available for realizing (2.3) and (2.4) in real time. Hence, the problem is concerned with determining A_e, B_e and C_e.

<u>Optimal Reduced-Order State-Estimation Problem (ROSE)</u>. Given the observed system

$$\dot{x} = (A+\sum_{i=1}^{p}\alpha_i A_i)x + w_1, \tag{2.6}$$

$$y = (C+\sum_{i=1}^{p}\alpha_i C_i) + w_2, \tag{2.7}$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \tag{2.8}$$

$$y_e = C_e x_e \tag{2.9}$$

which minimizes the state-estimation criterion

$$J(A_e,B_e,C_e) \triangleq \lim_{t\to\infty} \underline{\underline{E}}[(Lx-y_e)^T R(Lx-y_e)]. \tag{2.10}$$

To state the reduced-order dynamic-compensation problem, let x, A, $A_1, \ldots, A_p$, C, $C_1, \ldots, C_p$, y, w_1, w_2, V_1, V_{12}, V_2, and $\alpha_1, \ldots, \alpha_p$ be as in the reduced-order state-estimation problem. Furthermore, assume $u \in \underline{\underline{R}}^m$, $B \in \underline{\underline{R}}^{n\times m}$; R_1, R_{12} and R_2 are matrices of suitable dimension such that $R_1 \geq 0$, $R_2 \geq 0$ and $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$; $n_c \leq n$; and A_c, B_c and C_c are matrices of appropriate dimension. We require the technical assumption that, for each i, $B_i \neq 0$ implies $C_i = 0$. In this problem, the order n_c is considered fixed and is determined by implementation constraints. The dynamic compensator is assumed to have purely dynamic linear structure and the problem is concerned with determining A_c, B_c and C_c.

<u>Optimal Reduced-Order Dynamic-Compensation Problem (RODC)</u>. Given the control system

$$\dot{x} = (A+\sum_{i=1}^{p}\alpha_i A_i)x + (B+\sum_{i=1}^{p}\alpha_i B_i)u + w_1, \tag{2.11}$$

$$y = (C+\sum_{i=1}^{p}\alpha_i C_i)x + w_2 \tag{2.12}$$

design a reduced-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \tag{2.13}$$

$$u = C_c x_c \tag{2.14}$$

which minimizes the performance criterion

$$J(A_c,B_c,C_c) = \lim_{t\to\infty} \underline{\underline{E}}[x^T R_1 x + 2x^T R_{12} u + u^T R_2 u]. \tag{2.15}$$

Let n_r and (A_r,B_r,C_r) generically denote n_m, n_e, n_c and (A_m,B_m,C_m), (A_e,B_e,C_e), (A_c,B_c,C_c). To guarantee that J is finite and independent of initial conditions, we restrict (A_r,B_r,C_r) to the (open) set of second-moment-stable triples

$$\underline{S} \triangleq \left\{(A_r,B_r,C_r):\ \tilde{A}_s \otimes I_{n+n_r} + I_{n+n_r} \otimes \tilde{A}_s + \sum_{i=1}^{p} \tilde{A}_i \otimes \tilde{A}_i \text{ is stable}\right\},$$

where I_ν is the $\nu x \nu$ identity matrix, $\otimes$ denotes Kronecker product (see [120]),

$$\tilde{A}_s \triangleq \tilde{A} + \frac{1}{2}\sum_{i=1}^{p} \tilde{A}_i^2 ,$$

$$\tilde{A} \triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \quad \text{ROM};$$

$$\triangleq \begin{bmatrix} A & 0 \\ B_eC & A_e \end{bmatrix}, \quad \text{ROSE};$$

$$\triangleq \begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix}, \quad \text{RODC};$$

and

$$\tilde{A}_i \triangleq \begin{bmatrix} A_i & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{ROM};$$

$$\triangleq \begin{bmatrix} A_i & 0 \\ B_eC_i & 0 \end{bmatrix}, \quad \text{ROSE};$$

$$\triangleq \begin{bmatrix} A_i & B_iC_c \\ B_cC_i & 0 \end{bmatrix}, \quad \text{RODC}.$$

To avoid degeneracy in the derivation of the main results, we further restrict (without loss of generality) (A_r, B_r, C_r) to

$$\underline{S}_+ \triangleq \left\{(A_r, B_r, C_r) \in \underline{S}: \ (A_r, B_r) \text{ is controllable and } (A_r, C_r) \text{ is observable}\right\}.$$

Call a square matrix positive semisimple if it has positive eigenvalues and a diagonal Jordan canonical form, i.e., if it is similar to a positive diagonal (or, equivalently, a positive definite) matrix. Let $\rho(Z)$ denote the rank of a matrix Z. The following result is an immediate consequence of Theorem 6.2.5, p. 124 of [113].

Lemma 2.1. If nxn $\hat{Q}$, $\hat{P}$ are nonnegative definite and $\rho(\hat{Q}\hat{P}) = n_r$ then there exist $n_r$xn G, Γ and n_rxn_r positive-semisimple M such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \tag{2.16}$$

$$\Gamma G^T = I_{n_r}. \tag{2.17}$$

For convenience in stating the main results, we shall refer to $n_r \times n$ G, Γ and $n_r \times n_r$ positive-semisimple M satisfying (2.16) and (2.17) as a <u>(G,M,Γ)-factorization of $\hat{Q}\hat{P}$</u>. Also, we shall utilize the compact notation illustrated by

$$\underline{AQA}^T \triangleq \sum_{i=1}^{p} A_i Q A_i^T, \quad \underline{AQB} \triangleq \sum_{i=1}^{p} A_i Q B_i, \text{ etc.}$$

Hence, define for nonnegative-definite matrices Q, P, $\hat{Q}$ and $\hat{P}$ and a (G,M,Γ)-factorization of $\hat{Q}\hat{P}$,

$$A_s = A + \tfrac{1}{2}\underline{A}^2, \qquad B_s = B + \tfrac{1}{2}\underline{AB}, \qquad C_s = C + \tfrac{1}{2}\underline{CA},$$

$$R_{2s} \triangleq R_2 + \underline{B}^T(P+\hat{P})\underline{B}, \qquad V_{2s} \triangleq V_2 + \underline{C}(Q+\hat{Q})\underline{C}^T,$$

$$\underline{Q}_s \triangleq QC_s^T + V_{12} + \underline{A}(Q+\hat{Q})\underline{C}^T, \qquad \underline{P}_s \triangleq B_s^T P + R_{12}^T + \underline{B}^T(P+\hat{P})\underline{A},$$

$$\tau \triangleq G^T\Gamma, \quad \tau_1 \triangleq I_n - \tau.$$

<u>Theorem 2.1</u>. Suppose $(A_m, B_m, C_m) \in \underline{S}_+$ solves the optimal reduced-order modelling problem. Then there exist $n \times n$ nonnegative-definite matrices $\hat{Q}$ and $\hat{P}$ such that, for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$, A_m, B_m and C_m are given by

$$A_m = \Gamma A_s G^T, \tag{2.18}$$

$$B_m = \Gamma B, \tag{2.19}$$

$$C_m = CG^T, \tag{2.20}$$

and such that the following conditions are satisfied:

$$0 = A_s\hat{Q} + \hat{Q}A_s^T + BVB^T - \tau_\perp BVB^T\tau_\perp^T, \tag{2.21}$$

$$0 = A_s^T\hat{P} + \hat{P}A_s + C^TRC - \tau_\perp^T C^TRC\tau_\perp, \tag{2.22}$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_m. \tag{2.23}$$

<u>Theorem 2.2</u>. Suppose $(A_e, B_e, C_e) \in \underline{S}_+$ solves the optimal reduced-order state-estimation problem. Then there exist nxn nonnegative-definite matrices Q, $\hat{Q}$ and $\hat{P}$ such that, for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$, A_e, B_e and C_e are given by

$$A_e = \Gamma(A_s - \underline{Q}_s V_{2s}^{-1} C_s)G^T, \tag{2.24}$$

$$B_e = \Gamma \underline{Q}_s V_{2s}^{-1}, \tag{2.25}$$

$$C_e = LG^T, \tag{2.26}$$

and such that the following conditions are satisfied:

$$0 = A_sQ + QA_s^T + \underline{A}(Q+\hat{Q})\underline{A}^T + V_1 - \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T + \tau_\perp \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T\tau_\perp^T, \tag{2.27}$$

$$0 = A_s\hat{Q} + \hat{Q}A_s^T + \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T - \tau_\perp \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T\tau_\perp^T, \tag{2.28}$$

$$0 = (A_s - \underline{Q}_s V_{2s}^{-1}C_s)^T\hat{P} + \hat{P}(A_s - \underline{Q}_s V_{2s}^{-1}C_s) + L^TRL - \tau_\perp^T L^TRL\tau_\perp, \tag{2.29}$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_e. \tag{2.30}$$

<u>Theorem 2.3</u>. Suppose $(A_c, B_c, C_c) \in \underline{S}_+$ solves the optimal reduced-order dynamic-compensation problem. Then there exist nxn nonnegative-definite matrices Q, P, $\hat{Q}$ and $\hat{P}$ such that, for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$, A_c, B_c and C_c are given by

$$A_c = \Gamma(A_s - B_s R_{2s}^{-1}\underline{P}_s - \underline{Q}_s V_{2s}^{-1} C_s)G^T, \tag{2.31}$$

$$B_c = \Gamma \underline{Q}_s V_{2s}^{-1}, \tag{2.32}$$

$$C_c = -R_{2s}^{-1}\underline{P}_s G^T, \tag{2.33}$$

and such that the following conditions are satisfied:

$$0 = A_s Q + QA_s^T + \underline{A}Q\underline{A}^T + V_1 + (\underline{A}-\underline{B}R_{2s}^{-1}\underline{P}_s)\hat{Q}(\underline{A}-\underline{B}R_{2s}^{-1}\underline{P}_s)^T - \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T + \tau_\perp \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T \tau_\perp^T, \tag{2.34}$$

$$0 = A_s^T P + PA_s + \underline{A}^T P \underline{A} + R_1 + (\underline{A}-\underline{Q}_s V_{2s}^{-1}\underline{C})^T \hat{P}(\underline{A}-\underline{Q}_s V_{2s}^{-1}\underline{C}) - \underline{P}_s^T R_{2s}^{-1}\underline{P}_s + \tau_\perp^T \underline{P}_s^T R_{2s}^{-1}\underline{P}_s \tau_\perp, \tag{2.35}$$

$$0 = (A_s - B_s R_{2s}^{-1}\underline{P}_s)\hat{Q} + \hat{Q}(A_s - B_s R_{2s}^{-1}\underline{P}_s)^T + \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T - \tau_\perp \underline{Q}_s V_{2s}^{-1}\underline{Q}_s^T \tau_\perp^T, \tag{2.36}$$

$$0 = (A_s - \underline{Q}_s V_{2s}^{-1} C_s)^T \hat{P} + \hat{P}(A_s - \underline{Q}_s V_{2s}^{-1} C_s) + \underline{P}_s^T R_{2s}^{-1}\underline{P}_s - \tau_\perp^T \underline{P}_s^T R_{2s}^{-1}\underline{P}_s \tau_\perp, \tag{2.37}$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_c. \tag{2.38}$$

<u>Remark 2.1</u>. Because of (2.6) the nxn matrix τ which couples the design equations for each problem is idempotent, i.e., $\tau^2 = \tau$. In general this "optimal projection" is an <u>oblique</u> projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Note that from Sylvester's inequality it follows that $\rho(\tau) = n_r$.

<u>Remark 2.2</u>. Since $\hat{Q}\hat{P}$ is nonnegative semisimple it has a group generalized inverse $(\hat{Q}\hat{P})^\#$ given by $G^T M^{-1}\Gamma$ (see, e.g., [114], p. 124). Hence by (2.6) the optimal projection τ is given in closed form by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\#. \tag{2.39}$$

Remark 2.3. The Kalman filter and LQG results can be obtained as special cases of Theorems 2.2 and 2.3.

Remark 2.4. The following effects of the multiplicative white noise model are evident: (1) diagonal dominance in Q and P for modal systems and (2) suppressed system gains due to the definitions of V_{2s} and R_{2s}.

Remark 2.5. The formulation of the necessary conditions for feedback control in the presence of multiplicative white noise as a system of four matrix equations was discovered independently by Hyland [5] and Milshtein [89].

3. Design Capabilities and Limitations

In an attempt to minimize confusion in evaluating practicable design methodologies, we roughly categorize the pertinent issues according to (1) design tradeoffs, (2) modelling validity and (3) design-procedure complexity. The purpose of this section is to stress the multifaceted nature of the design process and to provide a setting in which to evaluate OP/ME design synthesis and alternative methods.

3.1 Design Tradeoffs

For control system design there are well-known tradeoffs among, for example, modelling validity, performance, stability, robustness, design-procedure complexity and implementation complexity. Although their relative importance is highly problem dependent, we feel that ultimately the most important attribute of a design methodology is not its ability to meet isolated design objectives but rater its ability to efficiently and optimally quantify tradeoffs among design objectives. Within the validity of its modelling formalism, optimal projection/maximum entropy design synthesis facilitates performance/robustness and performance/implementation complexity tradeoffs.

3.2 Modelling Validity

Stochastic linear ordinary differential equations with additive and multiplicative white disturbances permit representation of parameter uncertainties and, possibly mild nonlinearities. Treatment of nonwhite disturbances is possible by means of system augmentation, and extension to distributed parameter systems has been demonstrated ([21]). The quadratic performance criteria, although inappropriate for certain classes of problems, is particularly relevant to pointing and shaping of precision space structures. The state-space setting appears to permit more highly structured treatment of parameter errors than frequency domain approaches. The ramifications of the Stratonovich white-noise model, however, remain to be explored for general systems.

3.3 Design-Procedure Complexity

Evaluating advantages and disadvantages of a given design procedure is extremely complex because of dependence upon the underlying problem, design goals, design constraints, modelling validity and available design resources (human and computational). The underlying basis of the OP/ME design philosophy is to quantify available information, collocate available design variables and mechanize the design process to the greatest possible extent. This view is motivated by high-order systems with numerous strongly coupled sensors and actuators. Computational algorithms for solving the optimal projection equations have been developed in [14,17]. The principal difficulty arises from the presence of multiple extrema as a consequence of reduced-order design. As alluded to previously, however, uncertainty effects tend to simplify computations.

OPTIMAL PROJECTION DESIGN

DESIGN PROBLEM	DYNAMICAL SYSTEM	DESIGN OBJECTIVE	PERFORMANCE CRITERION	DESIGN EQUATIONS
REDUCED-ORDER MODELLING	$\dot{x} = Ax + Bw$ $y = Cx$	$\dot{x}_m = A_m x_m + B_m w$ $y_m = C_m x_m$	$\lim_{t\to\infty} E(y-y_m)^T R(y-y_m)$	2 Lyapunov
REDUCED-ORDER STATE ESTIMATION	$\dot{x} = Ax + w_1$ $y = Cx + w_2$	$\dot{x}_e = A_e x_e + B_e y$ $y_e = C_e x_e$	$\lim_{t\to\infty} E(Lx-y_e)^T R(Lx-y_e)$	1 Riccati 2 Lyapunov
REDUCED-ORDER DYNAMIC COMPENSATION	$\dot{x} = Ax + Bu + w_1$ $y = Cx + w_2$	$\dot{x}_c = A_c x_c + B_c y$ $u = C_c x_c$	$\lim_{t\to\infty} E(x^T R_1 x + u^T R_2 u)$	2 Riccati 2 Lyapunov

ALTERNATIVE SETTINGS

CONTINUOUS-TIME/DISCRETE-TIME

⇨ Digital implementation

INFINITE-DIMENSIONAL/FINITE-DIMENSIONAL

⇨ Distributed parameter systems

Figure 1

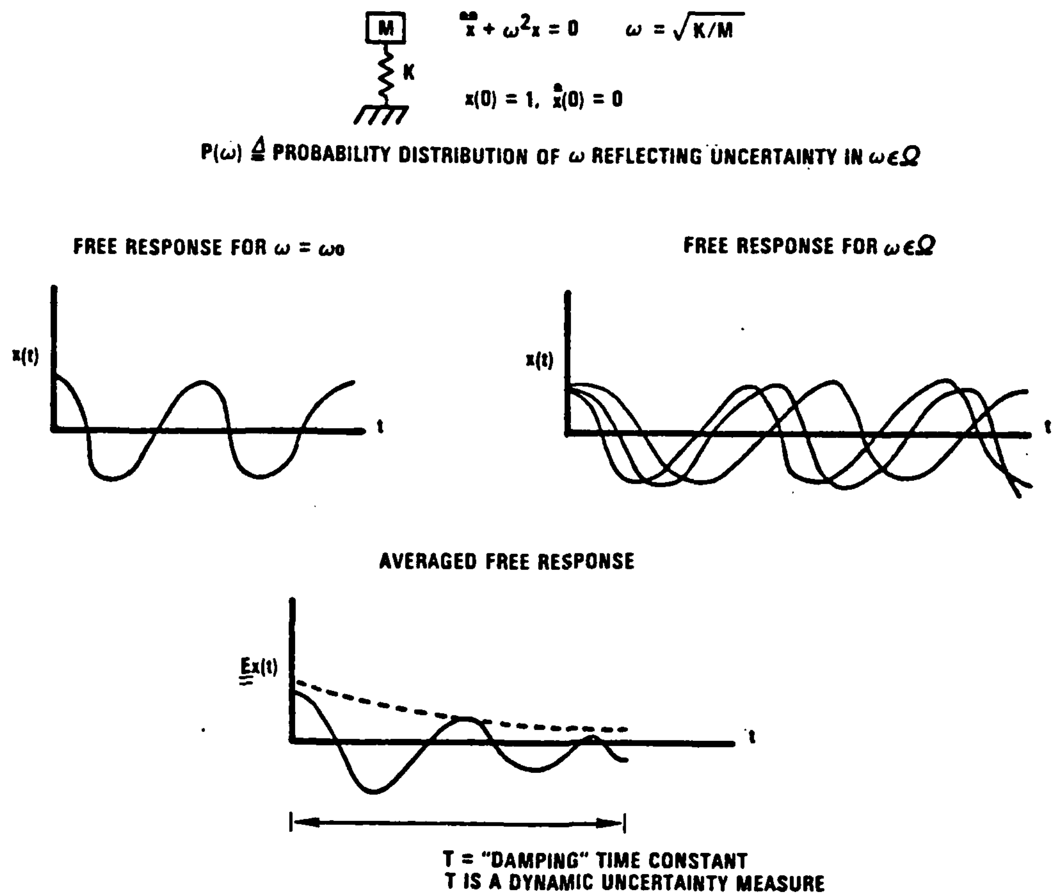
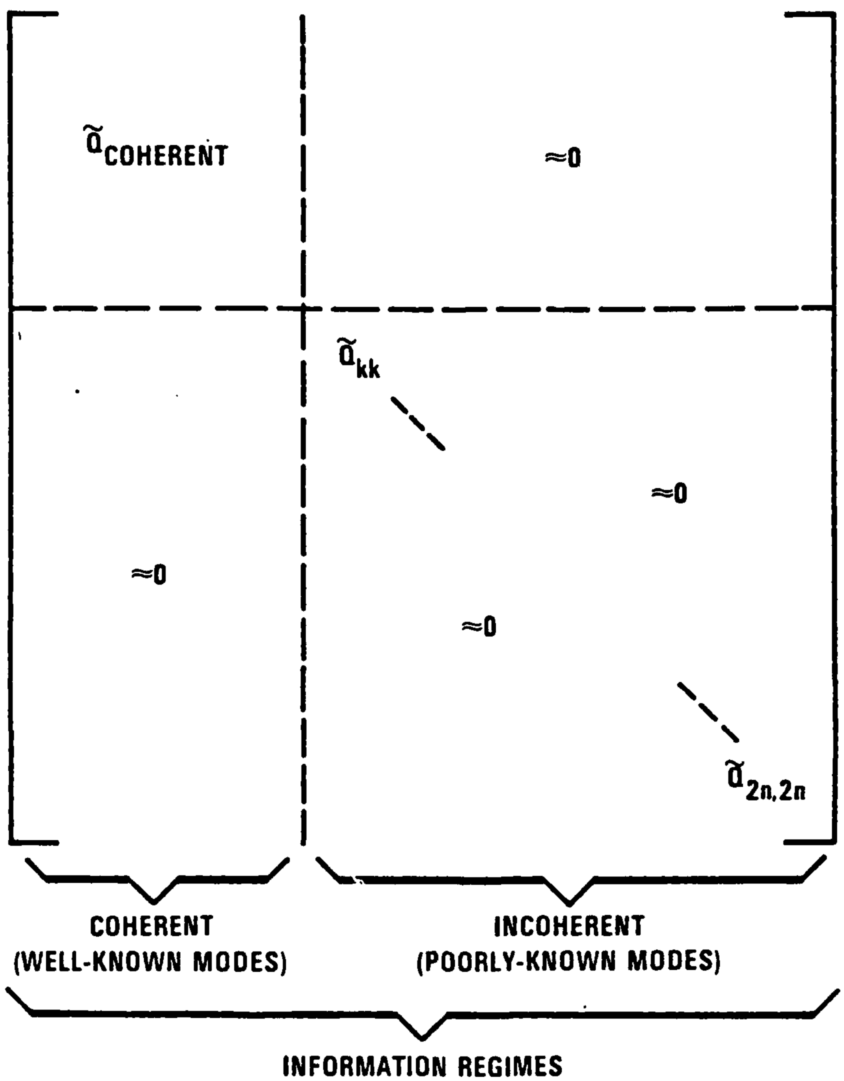


Figure 2. Decorrelation Time

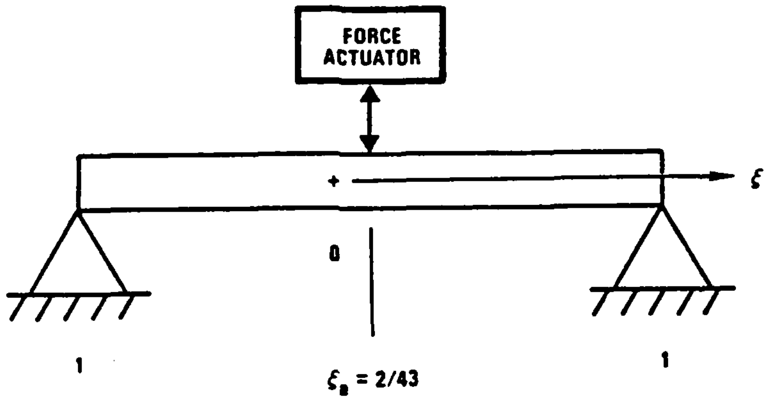


Qualitative Structure of Steady-State Covariance

$$\tilde{Q} \triangleq \lim_{t \to \infty} \tilde{Q}(t), \ \tilde{Q}(t) \triangleq \underline{\underline{E}}\left[\tilde{x}(t)\tilde{x}(t)^T\right]$$

(Frequency uncertainties increase with mode number)

Figure 3. Information Regimes



* NONDIMENSIONAL EQUATIONS OF MOTION ($\bar{\omega}_n = n^2$)
* "ENERGY" STATE-WEIGHTING
* UNCERTAINTIES IN OPEN-LOOP FREQUENCIES:

  $$T_K = \sqrt{\frac{\pi}{2}}(\sigma_K \bar{\omega}_K)^{-1}$$

  σ_K = STANDARD DEVIATION OF Kth MODE FREQUENCY
* SIMPLE UNCERTAINTY MODEL

  $$\sigma_K = \sigma \bar{\omega}_K$$

Figure 4. Simply-Supported Beam with Force Actuator Full-State Feedback Control Scheme

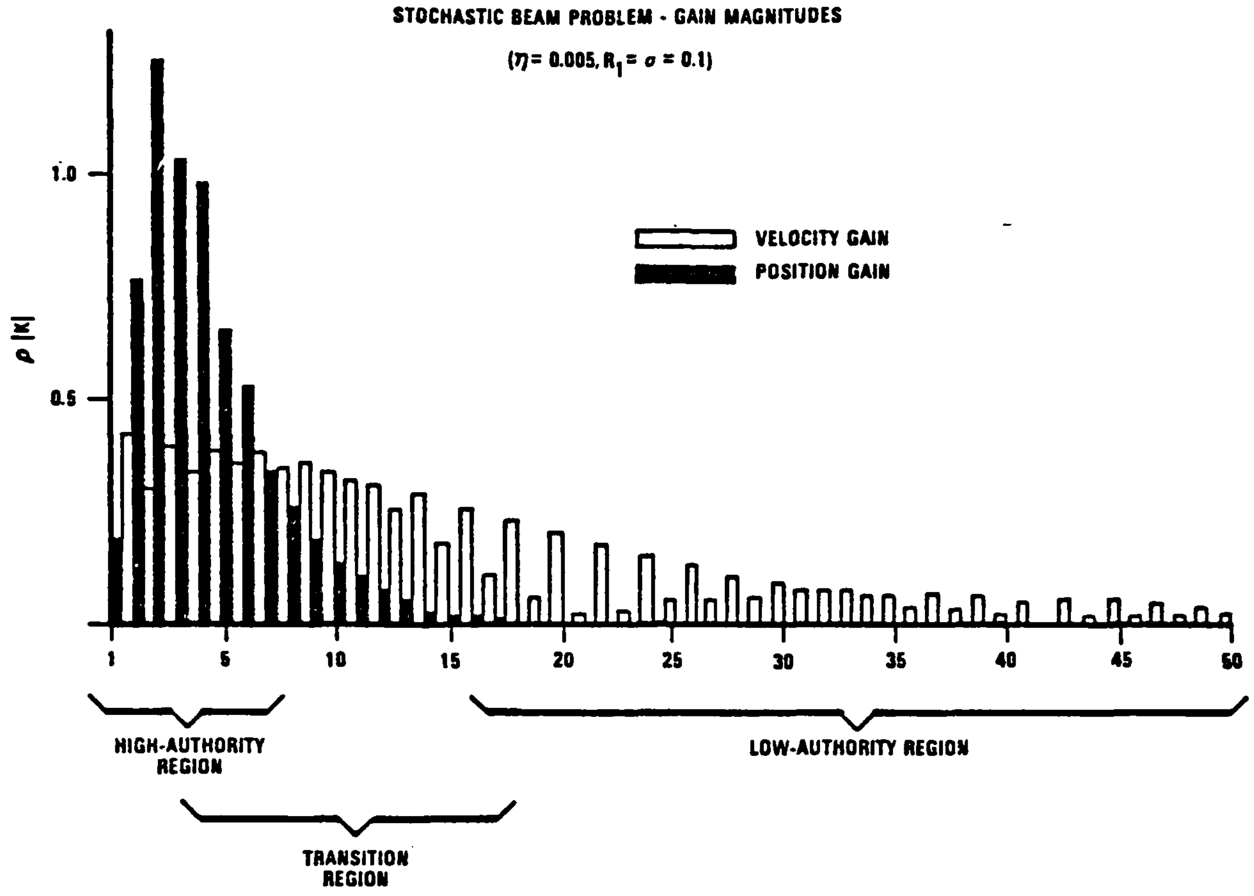


Figure 5. Authority Regions: A One-Step Direct Result of Maximum Entropy

References

1. D. C. Hyland, "Optimal Regulation of Structural Systems With Uncertain Parameters," MIT, Lincoln Laboratory, TR-551, 2 February 1981, DDC# AD-A099111/7.

2. D. C. Hyland, "Active Control of Large Flexible Spacecraft: A New Design Approach Based on Minimum Information Modelling of Parameter Uncertainties," VPI&SU/AIAA Symposium, Blacksburg, VA, June 1981.

3. D. C. Hyland, "Optimal Regulator Design Using Minimum Information Modelling of Parameter Uncertainties: Ramifications of the New Design Approach," VPI&SU/AIAA Symposium, Blacksburg, VA, June 1981.

4. D. C. Hyland and A. N. Madiwale, "Minimum Information Approach to Regulator Design: Numerical Methods and Illustrative Results," VPI&SU/AIAA Symposium, Blacksburg, VA, June 1981.

5. D. C. Hyland and A. N. Madiwale, "A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters," AIAA Guidance and Control Conference, Albuquerque, NM, August 1981.

6. D. C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters," AIAA 20th Aerospace Sciences Meeting, Orlando, FL, January 1982.

7. D. C. Hyland, "Structural Modeling and Control Design Under Incomplete Parameter Information: The Maximum Entropy Approach," AFOSR/NASA Workshops in Modeling, Analysis and Optimization Issues for Large Space Structures, Williamsburg, VA, May 1982.

8. D. C. Hyland, "Maximum Entropy Stochastic Approach to Control Design for Uncertain Structural Systems," American Control Conference, Arlington, VA, June 1982.

9. D. C. Hyland, "Minimum Information Stochastic Modelling of Linear Systems with a Class of Parameter Uncertainties," American Control Conference, Arlington, VA, June 1982.

10. D. C. Hyland and A. N. Madiwale, "Fixed-Order Dynamic Compensation Through Optimal Projection," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, JPL, Pasadena, CA, July 1982.

11. D. C. Hyland, "Minimum Information Modeling of Structural Systems with Uncertain Parameters," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, JPL, Pasadena, CA, July 1982.

12. D. C. Hyland, "Mean-Square Optimal Fixed-Order Compensation - Beyond Spillover Suppression," AIAA Astrodynamics Conference, San Diego, CA, August 1982.

13. D. C. Hyland, "Robust Spacecraft Control Design in the Presence of Sensor/Actuator Placement Errors," AIAA Astrodynamics Conference, San Diego, CA, August 1982.

14. D. C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results," AIAA 21st Aerospace Sciences Meeting, Reno, NV, January 1983.

15. D. C. Hyland, "Mean-Square Optimal, Full-Order Compensation of Structural Systems with Uncertain Parameters," MIT, Lincoln Laboratory TR-626, 1 June 1983.

16. D. C. Hyland and D. S. Bernstein, "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation," Proc. 22nd IEEE Conference on Decision and Control, San Antonio, TX, December 1983.

17. D. C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," Proc. AIAA Dynamics Specialists Conference, Palm Springs, CA, May 1984.

18. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation of Distributed Parameter Systems," Proc. AIAA Dynamics Specialists Conference, Palm Springs, CA, May 1984.

19. D. S. Bernstein and D. C. Hyland, "Numerical Solution of the Optimal Model Reduction Equations," AIAA Guidance and Control Conference, Seattle, WA, August 1984.

20. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore," 23rd IEEE Conference on Decision and Control, Las Vegas, NV, December 1984.

21. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Approach to Designing Optimal Finite-Dimensional Controllers for Distributed-Parameter Systems," 23rd IEEE Conference on Decision and Control, Las Vegas, NV, Dec. 1984.

22. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr. (to appear).

23. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," submitted for publication.

24. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," submitted for publication.

25. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," submitted for publication.

26. T. L. Johnson and M. Athans, "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 658-660, 1970.

27. W. S. Levine, T. L. Johnson and M. Athans, "Optimal Limited State Variable Feedback Controllers for Linear Systems," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 785-793, 1971.

28. K. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, New York, 1972.

29. D. B. Rom and P. E. Sarachik, "The Design of Optimal Compensators for Linear Constant Systems with Inaccessible States," IEEE Trans. Autom. Contr., Vol. AC-18, pp. 509-512, 1973.

30. M. Sidar and B.-Z. Kurtaran, "Optimal Low-Order Controllers for Linear Stochastic Systems," Int. J. Contr., Vol. 22, 377-387, 1975.

31. J. M. Mendel and J. Feather, "On the Design of Optimal Time-Invariant Compensators for Linear Stochastic Time-Invariant Systems," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 653-657, 1975.

32. S. Basuthakur and C. H. Knapp, "Optimal Constant Controllers for Stochastic Linear Systems," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 664-666, 1975.

33. R. B. Asher and J. C. Durrett, "Linear Discrete Stochastic Control with a Reduced-Order Dynamic Compensator," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 626-627, 1976.

34. W. J. Naeije and O. H. Bosgra, "The Design of Dynamic Compensators for Linear Multivariable Systems," 1977 IFAC, Fredricton, N.B., Canada, pp. 205-212.

35. H. R. Sirisena and S. S. Choi, "Design of Optimal Constrained Dynamic Compensators for Non-Stationary Linear Stochastic Systems," Int. J. Contr., Vol. 25, pp. 513-524, 1977.

36. P. J. Blanvillain and T. L. Johnson, "Invariants of Optimal Minimal-Order Observer Based Compensators," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 473-474, 1978.

37. C. J. Wenk and C. H. Knapp, "Parameter Optimization in Linear Systems with Arbitrarily Constrained Controller Structure," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 496-500, 1980.

38. J. O'Reilly, "Optimal Low-Order Feedback Controllers for Linear Discrete-Time Systems," in Control and Dynamic Systems, Vol. 16, edited C. T. Leondes, ed., Academic Press, 1980.

39. D. P. Looze and N. R. Sandell, Jr., "Gradient Calculations for Linear Quadratic Fixed Control Structure Problems," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 285-8, 1980.

40. D. A. Wilson, "Optimum Solution of Model-Reduction Problem," Proc. IEE, Vol. 117, pp. 1161-1165, 1970.

41. D. A. Wilson, "Model Reduction for Multivariable Systems," Int. J. Contr., Vol. 20, pp. 57-64, 1974.

42. R. N. Mishra and D. A. Wilson, "A New Algorithm for Optimal Reduction of Multivariable Systems," Int. J. Contr., Vol. 31, pp. 443-466, 1980.

43. D. A. Wilson and R. N. Mishra, "Design of Low Order Estimators Using Reduced Models," Int. J. Contr., Vol. 23, pp. 447-456, 1979.

44. C. S. Sims, "Reduced-Order Modelling and Filtering," in Control and Dynamic Systems, C. T. Leondes, ed., Vol. 18, pp. 55-103, 1982.

45. M. Aoki, "Control of Large-Scale Dynamic Systems by Aggregation," IEEE Trans. Auto. Contr., Vol. AC-13, pp. 246-253, 1968.

46. R. E. Skelton, "Cost Decomposition of Linear Systems with Application to Model Reduction," Int. J. Contr., Vol. 32, pp. 1031-1055, 1980.

47. B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 17-32, 1981.

48. L. Pernebo and L. M. Silverman, "Model Reduction via Balanced State Space Representations," IEEE Trans. Autom. Contr., Vol. AC-27, pp. 382-387, 1982.

49. K. V. Fernando and H. Nicholson, "On the Structure of Balanced and Other Principal Representations of SISO Systems," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 228-231, 1983.

50. S. Shokoohi, L. M. Silverman, and P. M. Van Dooren, "Linear Time-Variable Systems: Balancing and Model Reduction," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 810-822, 1983.

51. E. I. Verriest and T. Kailath, "On Generalized Balanced Realizations," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 833-844, 1983.

52. E. A. Jonckheere and L. M. Silverman, "A New Set of Invariants for Linear Systems - Application to Reduced-Order Compensator Design," IEEE Trans. Autom. Contr., Vol. AC-28, pp. 953-964, 1983.

53. R. E. Skelton and A. Yousuff, "Component Cost Analysis of Large Scale Systems," Int. J. Contr., Vol. 37, pp. 285-304, 1983.

54. A. Yousuff and R. E. Skelton, "Controller Reduction by Component Cost Analysis," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 520-530, 1984.

55. Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. Autom. Contr., Vol. AC-19, pp. 494-500, 1974.

56. Y. Bar-Shalom and E. Tse, "Generalized Certainty Equivalence and Dual Effect in Stochastic Control," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 817-819, 1975.

57. Y. Bar-Shalom and E. Tse, "Caution, Probing and the Value of Information in the Control of Uncertain Systems," Annals of Economic and Social Measurement, Vol. 5, pp. 323-337, 1976.

58. E. T. Jaynes, "New Engineering Applications of Information Theory," Proceedings of the First Symposium on Engineering applications of Random Function Theory and Probability, J. L. Bogdanoff and F. Kozin, pp. 163-203, Wiley, New York, 1963.

59. E. T. Jaynes, "Prior Probabilities," IEEE Trans. Sys. Sci. Cybern., Vol. SSC-4, pp. 227-241, 1968.

60. E. T. Jaynes, "Where Do We Stand on Maximum Entropy?," The Maximum Entropy Formalism, D. Levine and M. Tribus, eds., The MIT Press, pp. 15-118, Cambridge, MA, 1979.

61. R. D. Rosenkrandtz, ed., "E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics," Reidel, Boston, 1983.

62. K. Ito, On Stochastic Differential Equations, Amer. Math. Soc., Providence, RI, 1951.

63. E. Wong and M. Zakai, "On the Relation Between Ordinary and Stochastic Differential Equations," Int. J. Engrg. Sci., Vol. 3, pp. 213-229, 1965.

64. R. L. Stratonovich, "A New Representation for Stochastic Integrals," SIAM J. Contr., Vol. 4, pp. 362-371, 1966.

65. R. L. Stratonovich, Conditional Markov Process and Their Application to the Theory of Optimal Control, Elsevier, NY, 1968.

66. A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

67. E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

68. E. J. McShane, Stochastic Calculus and Stochastic Models, Academic Press, New York, 1974.

69. L. Arnold, Stochastic Differential Equations: Theory and Applications, Wiley, New York, 1974.

70. W. H. Fleming and R. W. Rishel, Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 1975.

71. H. J. Sussmann, "On the Gap Between Deterministic and Stochastic Ordinary Differential Equations," The Annals of Probability, Vol. 6, pp. 19-41, 1978.

72. W. M. Wonham, "Optimal Stationary Control of Linear Systems with State-Dependent Noise," SIAM J Contr., Vol. 5, pp. 486-500, 1967.

73. M. Metivier and J. Pellaumail, Stochastic Integration, Academic Press, New York, 1980.

74. W. M. Wonham, "On a Matrix Riccati Equation of Stochastic Control," SIAM J. Contr., Vol. 6, pp. 681-697, 1968.

75. W. M. Wonham, "Random Differential Equations in Control Theory," in Probabilistic Analysis in Applied Mathematics, A. T. Bharucha-Reid, ed., Vol. 2, pp. 131-212, Academic Press, New York, 1970.

76. D. Kleinman, "Optimal Stationary Control of Linear Systems with Control-Dependent Noise," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 673-677, 1969.

77. P. J. McLane, "Optimal Linear Filtering for Linear Systems with State-Dependent Noise," Int. J. Contr., Vol. 10, pp. 41-51, 1969.

78. P. McLane, "Optimal Stochastic Control of Linear Systems with State- and Control-Dependent Disturbances," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 793-798, 1971.

79. D. Kleinman, "Numerical Solution of the State Dependent Noise Problem," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 419-420, 1976.

80. U. Haussmann, "Optimal Stationary Control with State and Control Dependent Noise," SIAM J. Contr., Vol. 9, pp. 184-198, 1971.

81. J. Bismut, "Linear-Quadratic Optimal Stochastic Control with Random Coefficients," SIAM J. Contr., Vol. 14, pp. 419-444, 1976.

82. A. Ichikawa, "Optimal Control of a Linear Stochastic Evolution Equation with State and Control Dependent Noise," Proc. IMA Conference on Recent Theoretical Development in Control, Leicester, England, Academic Press, New York, 1976.

83. A. Ichikawa, "Dynamic Programming Approach to Stochastic Evolution Equations," SIAM J. Contr. Optim., Vol 17, pp. 152-174, 1979.

84. N. U. Ahmed, "Stochastic Control on Hilbert Space for Linear Evolution Equations with Random Operator-Valued Coefficients," SIAM J. Contr. Optim., Vol. 19, pp. 401-430.

85. C. W. Merriam III, Automated Design of Control Systems, Gordon and Breach, New York, 1974.

86. M. Aoki, "Control of Linear Discrete-Time Stochastic Dynamic Systems with Multiplicative Disturbances," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 388-392, 1975.

87. D. E. Gustafson and J. L. Speyer, "Design of Linear Regulators for Nonlinear Systems," J. Spacecraft and Rockets, Vol. 12, pp. 351-358, 1975.

88. D. E. Gustafson and J. L. Speyer, "Linear Minimum Variance Filters Applied to Carrier Tracking," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 65-73, 1976.

89. G. N. Milshtein, "Design of Stabilizing Controller with Incomplete State Data for Linear Stochastic System with Multiplicative Noise," Autom. and Remote Contr., Vol. 43, pp. 653-659, 1982.

90. M. Athans, R. T. Ku and S. B. Gershwin, "The Uncertainty Threshold Principle: Some Fundamental Limitations of Optimal Decision Making Under Dynamic Uncertainty," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 491-495, 1977.

91. R. J. Ku and M. Athans, "Further Results on the Uncertainty Threshold Principle," IEEE Trans. Autom. Contr., Vol. AC-22, pp. 866-868, 1977.

92. F. Kozin, "A Survey of Stability of Stochastic Systems," Automatica, Vol. 5, pp. 95-112, 1969.

93. D. L. Kleinman, "On the Stability of Linear Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-14, pp. 429-430, 1969.

94. U. G. Haussmann, "Stability of Linear Systems with Control Dependent Noise," SIAM J. Contr., Vol. 11, pp. 382-394, 1973.

95. A. S. Willsky, S. I. Marcus and D. N. Martin, "On the Stochastic Stability of Linear Systems Containing Colored Multiplicative Noise," IEEE Trans. Autom. Contr., Vol. AC-20, pp. 711-713, 1975.

96. J. C. Willems and G. C. Blankenship, "Frequency Domain Stability Criteria for Stochastic Systems," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 292-299, 1971.

97. J. L. Willems, "Mean Square Stability Criteria for Stochastic Feedback Systems," Int. J. Sys. Sci., Vol. 4, pp. 545-564, 1973.

98. U. Haussmann, "On the Existence of Moments of Stationary Linear Systems with Multiplicative Noise," SIAM J. Contr., Vol. 12, pp. 99-105, 1974.

99. J. L. Willems and J. C. Willems, "Feedback Stabilizability for Stochastic Systems with State and Control Dependent Noise," Automatica, Vol. 12, pp. 277-283, 1976.

100. J. L. Willems, "Moment Stability of Linear White Noise and Coloured Noise Systems," in Stochastic Problems in Dynamics, pp. 36-53, B. L. Clarkson, ed., Pitman, London, 1977.

101. A. Kistner, "On the Moments of Linear Systems Excited by a Coloured Noise Process," in Stochastic Problems in Dynamics, pp. 36-53, B. L. Clarkson, ed., Pitman, London, 1977.

102. T. Sasagawa, "On the Exponential Stability and Instability of Linear Stochastic Systems," Int. J. Contr., Vol. 33, pp. 363-370, 1981.

103. T. Sasagawa, "Sufficient Conditions for the Exponential p-Stability and p-Stabilizability of Linear Stochastic Systems," Int. J. Sys. Sci., Vol. 13, pp. 399-408, 1982.

104. Y. A. Phillis, "Entropy Stability of Continuous Dynamic Systems," Int. J. Control, Vol. 35, pp. 323-340, 1982.

105. Y. A. Phillis, "Optimal Stabilization of Stochastic Systems," J. Math. Anal. Applic., Vol. 94, pp. 489-500, 1983.

106. J. L. Willems and J. C. Willems, "Robust Stabilization of Uncertain Systems," SIAM J. Contr. Optim., Vol. 21, pp. 352-374, 1983.

107. F. M. Brasch and J. B. Pearson," Pole Placement Using Dynamic Compensators," IEEE Trans. Autom. Contr., Vol. AC-15, pp. 34-43, 1970.

108. R. Ahmari and A. C. Vacroux, "On the Pole Assignment in Linear Systems with Fixed-Order Compensators," Int. J. Contr., Vol. 17, pp. 397-404, 1973.

109. D. C. Youla, J. J. Bongiorno, Jr., and C. N. Lu, "Single-Loop Feedback-Stabilization of Linear Multivariable Dynamical Plants," Automatica, Vol. 10, pp. 159-173, 1974.

110. H. Seraji, "An Approach to Dynamic Compensator Design for Pole Assignment," Int. J. Contr., Vol. 21, pp. 955-966, 1975.

111. R. V. Patel, "Design of Dynamic Compensators for Pole Assignment," Int. J. Systems Sci., Vol. 7, pp. 207-224, 1976.

112. C. A. Desoer, R. W. Liu, J. Murray, and R. Saeks, "Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 399-412, 1980.

113. C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Its Applications, John Wiley and Sons, New York, 1971.

114. S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

115. M. Athans, "The Matrix Minimum Principle," Inform. Contr., Vol. 11, pp. 592-606, 1968.

116. W. M. Wonham, Linear Multivariable Control: A Geometric Approach, Springer-Verlag, New York, 1974.

117. T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.

118. A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse," SIAM J. Appl. Math., Vol. 17, pp. 434-440, 1969.

119. E. Kreindler and A. Jameson, "Conditions for Nonnegativeness of Partitioned Matrices," IEEE Trans. Autom. Contr., Vol. AC-17, pp. 147-8, 1972.

120. S. Barnett and C. Storey, Matrix Methods in Stability Theory, Barnes and Noble, New York, 1976.

121. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," IEEE Trans. Autom. Contr., Vol AC-17, pp. 474-483, 1972.

122. A. Vinkler and L. J. Wood, "Multistep Guaranteed Cost control of Linear Systems with Uncertain Parameters," J. Guidance and Control, Vol. 2, pp. 449-456, 1979.

APPENDIX F

(REFERENCE [30])

THE OPTIMAL PROJECTION EQUATIONS FOR REDUCED-ORDER STATE ESTIMATION

Dennis S. Bernstein
Harris Corp., GASD
P. O. Box 37
Melbourne, Florida 32901

David C. Hyland
Harris Corp., GASD
P. O. Box 37
Melbourne, Florida 32901

Abstract

First-order necessary conditions for optimal, steady-state, reduced-order state estimation for a linear, time-invariant plant in the presence of correlated disturbance and nonsingular *measurement* noise are derived in a new and highly-simplified form. In contrast to the lone matrix Riccati equation arising in the full-order (Kalman filter) case, the optimal steady-state reduced-order estimator is characterized by three matrix equations (one modified Riccati equation and two modified Lyapunov equations) coupled by a projection whose rank is precisely equal to the order of the estimator and which determines the optimal estimator gains. This coupling is a graphic reminder of the suboptimality of proposed approaches involving either model reduction followed by "full-order" estimator design or full-order estimator design followed by estimator-reduction techniques. The results given here complement recently-obtained results which characterize the optimal reduced-order model by means of a pair of coupled modified Lyapunov equations ([7]) and the optimal fixed-order dynamic compensator by means of a coupled system of two modified Riccati equations and two modified Lyapunov equations ([6]).

1. Introduction

It has recently been shown (see [1-7]) that the first-order necessary conditions for the problems of optimal model reduction and optimal fixed-order dynamic compensation can be formulated in terms of an "optimal projection" matrix which arises as a direct consequence of optimality. These necessary conditions, by virtue of their remarkable simplicity, yield insight into the structure of the optimal design and permit the development of alternative numerical algorithms ([2,4,7]). The purpose of this note is to develop analogous first-order necessary conditions for the reduced-order state-estimation problem. Since this problem falls midway between the problems of open-loop model reduction and closed-loop fixed-order dynamic compensation, it is not surprising that the necessary conditions for these problems are correspondingly related. Specifically, while the optimal projection equations for model reduction consist of a system of <u>two</u> matrix equations (a pair of modified Lyapunov equations) and the optimal projection equations for fixed-order dynamic compensation comprise a system of <u>four</u> matrix equations (a pair of modified Lyapunov equations plus a pair of modified Riccati equations), the optimal projection equations for reduced-order state estimation form a system of <u>three</u> matrix equations (a pair of modified Lyapunov equations along with a single modified Riccati equation). In each case the system of matrix equations is coupled by an <u>oblique</u> projection (idempotent matrix) which determines the gains of the optimal reduced-order system, whether it be a model, estimator or compensator.

The need for designing an optimal reduced-order state estimator for a high-order dynamic system follows directly from real-world constraints on computing capability. A further motivation is the fact that although a system may have many degrees of freedom, it is often the case that estimates of only a small number of state variables are actually required. In the face of these practical motivations, numerous approaches to designing reduced-order state estimators have been proposed. See [8] for a recent review of previous results.

An important fact pointed out in [8] and [9] is that reduced-order estimators designed by means of either model reduction followed by "full-order" state estimation or full-order estimation followed by estimator reduction will not be optimal for the given order. In the present paper this point is graphically confirmed by the fact that the three matrix equations characterizing the optimal reduced-order state estimator reveal intrinsic coupling (via the optimal projection) between the "operations" of optimal estimation (the modified Riccati equation) and optimal model reduction (the pair of modified Lyapunov equations).

2. Problem Statement and Main Result

The following notation and definitions will be used throughout the paper:

n, ℓ, n_e, p	positive integers, $1 \le n_e \le n$
x, y, x_e, y_e	n, ℓ, n_e, p-dimensional vectors
A, C, L	$n\times n$, $\ell\times n$, $p\times n$ matrices
A_e, B_e, C_e	$n_e\times n_e$, $n_e\times\ell$, $p\times n_e$ matrices
$w_1(t)$, $t \ge 0$	n-dimensional white noise with nonnegative-definite intensity V_1
$w_2(t)$, $t \ge 0$	ℓ-dimensional white noise with positive-definite intensity V_2
V_{12}	$n\times\ell$ matrix satisfying $\mathbb{E}[w_1(t)w_2(s)^T] = V_{12}\delta(t-s)$
R	$p\times p$ positive-definite matrix
I_r	$r\times r$ identity matrix
Z^T	transpose of vector or matrix Z
Z^{-T}	$(Z^T)^{-1}$ or $(Z^{-1})^T$
$\mathcal{N}(Z)$, $\mathcal{R}(Z)$, $\rho(Z)$	null space, range, rank of matrix Z
$\mathbb{E}$	expected value
$\mathbb{R}$, $\mathbb{R}^{r\times s}$	real numbers, $r\times s$ real matrices
stable matrix	matrix with eigenvalues in open left half plane
nonnegative-definite matrix	symmetric matrix with nonnegative eigenvalues
positive-definite matrix	symmetric matrix with positive eigenvalues
nonnegative-semisimple matrix	matrix similar to a nonnegative-definite matrix
positive-semisimple matrix	matrix similar to a positive-definite matrix
positive-diagonal matrix	diagonal matrix with positive diagonal elements

We consider the following optimal reduced-order state-estimation problem. Given the system

$$\dot{x} = Ax + w_1, \tag{2.1}$$

$$y = Cx + w_2, \tag{2.2}$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \tag{2.3}$$

$$y_e = C_e x_e, \tag{2.4}$$

which minimizes the error criterion

$$J(A_e,B_e,C_e) \triangleq \lim_{t\to\infty} \mathbb{E}[(Lx-y_e)^T R(Lx-y_e)].$$

In this formulation the matrix L identifies the states, or linear combination of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (2.3), (2.4) in real time. Hence, n_e is considered to be fixed in what follows and the problem is concerned with determining A_e, B_e and C_e.

To guarantee that J is finite it is assumed that A is stable and we restrict our attention to the set of stable reduced-order estimators

$$\mathcal{A} \triangleq \{(A_e,B_e,C_e): A_e \text{ is stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (2.3) and (2.4), without loss of generality we further restrict our attention to the set of admissible estimators

$$\mathcal{A}_+ \triangleq \{A_e,B_e,C_e) \in \mathcal{A}: (A_e,B_e) \text{ is controllable and } (A_e,C_e) \text{ is observable}\}.$$

The following lemma, whose proof is given in [7], is needed for the statement of the main result.

Lemma 2.1. Suppose $\hat{Q},\hat{P} \in \mathbb{R}^{n\times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_e$ then there exist $G,\Gamma \in \mathbb{R}^{n_e\times n}$ and positive-semisimple $M \in \mathbb{R}^{n_e\times n_e}$ such that

$$\hat{Q}\hat{P} = G^T M\Gamma, \tag{2.5}$$

$$\Gamma G^T = I_{n_e}. \tag{2.6}$$

For convenience in stating the Main Theorem we shall refer to $G,\Gamma \in \mathbb{R}^{n_e\times n}$ and positive-semisimple $M \in \mathbb{R}^{n_e\times n_e}$ satisfying (2.5) and (2.6) as a (G,M,Γ)-factorization of $\hat{Q}\hat{P}$. Furthermore, define the notation

$$\tau \triangleq G^T\Gamma, \quad \tau_\perp \triangleq I_n - \tau$$

and

$$\mathcal{Q} \triangleq QC^T + V_{12},$$

where $Q \in \mathbb{R}^{n\times n}$.

Main Theorem. Suppose $(A_e,B_e,C_e) \in \mathcal{A}_+$ solves the optimal reduced-order state-estimation problem. Then there exist nonnegative-definite matrices Q, $\hat{Q}$, $\hat{P} \in \mathbb{R}^{n\times n}$ such that, for some (G,M,Γ)-factorization of $\hat{Q}\hat{P}$, A_e, B_e and C_e are given by

$$A_e = \Gamma(A-\mathcal{Q}V_2^{-1}C)G^T, \tag{2.7}$$

$$B_e = \Gamma\mathcal{Q}V_2^{-1}, \tag{2.8}$$

$$C_e = LG^T, \tag{2.9}$$

and such that the following conditions are satisfied:

$$0 = AQ + QA^T + V_1 - QV_2^{-1}Q^T + \tau_\perp QV_2^{-1}Q^T\tau_\perp^T, \tag{2.10}$$

$$0 = A\hat{Q} + \hat{Q}A^T + QV_2^{-1}Q^T - \tau_\perp QV_2^{-1}Q^T\tau_\perp^T, \tag{2.11}$$

$$0 = (A-QV_2^{-1}C)^T\hat{P} + \hat{P}(A-QV_2^{-1}C) + L^TRL - \tau_\perp^T L^TRL\tau_\perp, \tag{2.12}$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_e. \tag{2.13}$$

<u>Remark 2.1</u>. It is useful to note that (2.7) can be replaced by

$$A_e = \Gamma AG^T - B_eCG^T. \qquad (2.7)'$$

<u>Remark 2.2</u>. Because of (2.6) the $n\times n$ matrix τ which couples the three equations (2.10)-(2.12) is idempotent, i.e., $\tau^2 = \tau$. In general this "optimal projection" is an <u>oblique</u> projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Note that from Sylvester's inequality and (2.6) it follows that $\rho(\tau) = n_e$. It should be stressed that the form of the optimal reduced-order estimator (2.7)-(2.9) is a direct consequence of optimality and not the result of an a priori assumption on the structure of the reduced-order estimator.

<u>Remark 2.3</u>. To obtain the standard steady-state Kalman filter result for the full-order case, set $p = n_e = n$ and $L = I_n$. Then $\tau = G = \Gamma = I_n$ and thus (2.10) reduces to the standard observer Riccati equation ([10], p. 367) and (2.7) and (2.8) yield the usual expressions. Furthermore, it follows from (2.7)', Lemma 2.1 of [11] and standard results that (2.11)-(2.13) are equivalent to the assumption that (A_e, B_e, C_e) is controllable and observable.

<u>Remark 2.4</u>. Since $\hat{Q}\hat{P}$ is nonnegative semisimple it has a group generalized inverse $(\hat{Q}\hat{P})^{\#}$ given by $G^TM^{-1}\Gamma$ (see, e.g., [12], p. 124). Hence by (2.6) the optimal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}. \tag{2.14}$$

Remark 2.5. Replacing x_e by Sx_e, where S is invertible, yields the "equivalent" estimator $(SA_eS^{-1}, SB_e, C_eS^{-1})$. Since $J(A_e,B_e,C_e) = J(SA_eS^{-1},SB_e,C_eS^{-1})$ one would expect the Main Theorem to apply also to $(SA_eS^{-1},SB_e,C_eS^{-1})$. This is indeed the case since transformation of the estimator state basis corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T(SMS^{-1})(S\Gamma)$.

Remark 2.6. Note that, for the optimal values of A_e, B_e and C_e, (2.3) assumes the observer form

$$\dot{x}_e = \Gamma AG^T x_e + \Gamma \mathcal{Q} V_2^{-1}(y - CG^T x_e). \tag{2.15}$$

By introducing the quasi-full-state estimate $\hat{x} \triangleq G^T x_e \in \mathbb{R}^n$ so that $\tau\hat{x} = \hat{x}$ and $x_e = \Gamma\hat{x} \in \mathbb{R}^{n_e}$, (2.15) can be written as

$$\dot{\hat{x}} = \tau A\tau\hat{x} + \tau\mathcal{Q}V_2^{-1}(y - C\hat{x}). \tag{2.16}$$

Note that although the implemented estimator (2.15) has the state $x_e \in \mathbb{R}^{n_e}$, (2.15) can be viewed as a quasi-full-order estimator whose geometric structure is entirely dictated by the projection τ. Specifically, error inputs $\mathcal{Q}V_2^{-1}(y - C\hat{x})$ are annihilated unless they are contained in $[\mathcal{N}(\tau)]^{\perp} = \mathcal{R}(\tau^T)$. Hence the observation subspace of the estimator is precisely $\mathcal{R}(\tau^T)$.

Remark 2.7. Although the form of (2.16) would lead one to surmise that the optimal reduced-order estimator is a projection of the optimal full-order estimator, this is not generally the case for the following simple reason. In the full-order case Q (which appears in $\mathcal{Q}$) is determined by solving a single Riccati equation whereas in the reduced-order case Q must be found in conjunction with $\hat{Q}$ and $\hat{P}$ to satisfy all three matrix equations (2.10)-(2.12). Hence the value of Q in the reduced-order case may be different from the value of Q in the full-order case. Thus (2.16) may not be obtainable by simply projecting the full-order result.

To further clarify the relationship between $\hat{Q}$, $\hat{P}$ and τ, we now show that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ.

<u>Proposition 2.1</u>. There exists invertible $\Phi \in \mathbb{R}^{n\times n}$ such that

$$\hat{Q} = \Phi^{-1}\begin{bmatrix}\Lambda_{\hat{Q}} & 0\\ 0 & 0\end{bmatrix}\Phi^{-T}, \qquad \hat{P} = \Phi^{T}\begin{bmatrix}\Lambda_{\hat{P}} & 0\\ 0 & 0\end{bmatrix}\Phi, \tag{2.17}$$

$$\hat{Q}\hat{P} = \Phi^{-1}\begin{bmatrix}\Lambda & 0\\ 0 & 0\end{bmatrix}\Phi, \qquad \tau = \Phi^{-1}\begin{bmatrix}I_{n_e} & 0\\ 0 & 0\end{bmatrix}\Phi, \tag{2.18a,b}$$

where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \mathbb{R}^{n_e\times n_e}$ are positive diagonal, $\Lambda \triangleq \Lambda_{\hat{Q}}\Lambda_{\hat{P}}$ and the diagonal elements of Λ are the eigenvalues of M. Consequently,

$$\hat{Q} = \tau\hat{Q}, \qquad \hat{P} = \hat{P}\tau. \tag{2.19}$$

3. Proof of the Main Theorem

The proof proceeds exactly as in [6]. Using the fact that $\mathcal{A}_+$ is open, the Fritz John version of the Lagrange multiplier theorem can be used to rigorously derive the first-order necessary conditions

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}, \tag{3.1}$$

$$0 = \tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{R}, \tag{3.2}$$

$$0 = P_{12}^T Q_{12} + P_2 Q_2, \tag{3.3}$$

$$B_e = -[(P_2^{-1}P_{12}^T Q_1 + Q_{12}^T)C^T + P_2^{-1}P_{12}^T V_{12}]V_2^{-1}, \tag{3.4}$$

$$C_e = SQ_{12}Q_2^{-1}, \tag{3.5}$$

where

$$\tilde{A} = \begin{bmatrix}A & 0\\ B_eC & A_e\end{bmatrix}, \quad \tilde{V} = \begin{bmatrix}V_1 & V_{12}B_e^T\\ B_eV_{12} & B_eV_2B_e^T\end{bmatrix}, \quad \tilde{R} = \begin{bmatrix}S^TRS & -S^TRC_e\\ -C_e^TRS & C_e^TR_2C_e\end{bmatrix}$$

and $(n+n_e)x(n+n_e)$ $\tilde{Q}$, $\tilde{P}$ are partitioned into nxn, nxn_e and $n_e x n_e$ subblocks as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Expanding (3.1) and (3.2) yields

$$0 = AQ_1 + Q_1A^T + V_1, \tag{3.6}$$

$$0 = AQ_{12} + Q_{12}A_e^T + Q_1(B_eC)^T + V_{12}B_e^T, \tag{3.7}$$

$$0 = A_eQ_2 + Q_2A_e^T + B_eCQ_{12} + Q_{12}^T(B_eC)^T + B_eV_2B_e^T, \tag{3.8}$$

$$0 = A^TP_1 + P_1A + (B_eC)^TP_{12}^T + P_{12}B_eC + S^TRS, \tag{3.9}$$

$$0 = P_{12}A_e + A^TP_{12} + (B_eC)^TP_2 - S^TRC_e, \tag{3.10}$$

$$0 = A_e^TP_2 + P_2A_e + C_e^TR_2C_e. \tag{3.11}$$

Note that (3.9) is superfluous and can be omitted. Writing (3.8) as (see [13,14])

$$0 = (A_e+B_eCQ_{12}Q_2^+)Q_2 + Q_2(A_e+B_eCQ_{12}Q_2^+)^T + B_eV_2B_e^T,$$

where Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2, it follows from Lemmas 2.1 and 12.2 of [11] that Q_2 is positive definite. Similarly, (3.11) implies that P_2 is positive definite. This justifies (3.4) and (3.5).

Now define the nxn nonnegative-definite matrices (see [13,14])

$$Q = Q_1-Q_{12}Q_2^{-1}Q_{12}^T, \quad \hat{Q} = Q_{12}Q_2^{-1}Q_{12}^T, \quad \hat{P} = P_{12}P_2^{-1}P_{12}^T,$$

and note that (3.3) implies (2.5) and (2.6) with

$$G = Q_2^{-1}Q_{12}^T, \quad M = Q_2P_2, \quad \Gamma = -P_2^{-1}P_{12}^T .$$

Since $Q_2P_2 = P_2^{-1/2}(P_2^{1/2}Q_2P_2^{1/2})P_2^{1/2}$, M is positive semisimple. Sylvester's inequality yields (2.13). Note (2.19) and the identities

$$Q_1 = Q + \hat{Q}, \tag{3.12}$$

$$Q_{12} = \hat{Q}\Gamma^T, \; P_{12} = -\hat{P}G^T, \tag{3.13}$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \; P_2 = G\hat{P}G^T. \tag{3.14}$$

Using (3.12)-(3.14), (3.4) and (3.5) yield (2.8) and (2.9). Also, RHS(3.8)-RHS(3.7) yields (2.7). Substituting (2.7)-(2.9) into (3.6)-(3.8), (3.10) and (3.11), it can be seen that (3.8) and (3.11) are also superfluous. Finally, linear combinations of the remaining three equations (3.6), (3.7) and (3.10) yield (2.10)-(2.12).

4. Concluding Remarks

The question of multiple local minima satisfying the optimal projection equations for reduced-order state estimation and the problem of constructing numerical methods for solving these equations are beyond the scope of this note. It should be pointed out, however, that promising numerical results for the model-reduction and fixed-order dynamic-compensation problems have been obtained by means of iterative algorithms that take full advantage of the presence and structure of the optimal projection ([2,4,7]).

Finally, the results of this paper can be extended to include the following related problems: 1) discrete-time system/discrete-time estimator; 2) infinite-dimensional system/finite-dimensional estimator ([5]); and 3) parameter uncertainties ([1,15]).

Acknowledgment. We wish to thank Dr. Fredric M. Ham for directing our attention to the reduced-order state-estimation problem as a fruitful application of the optimal projection approach.

References

1. D.C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters", AIAA 20th Aerospace Sciences Mtg., Orlando, FL, Jan. 1982.

2. D.C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results", AIAA 21st Aerospace Sciences Mtg., Reno, NV, Jan. 1983.

3. D.C. Hyland and D.S. Bernstein, "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation", Proc. 22nd IEEE Conf. on Decision and Control, San Antonio, TX, Dec. 1983.

4. D.C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection", AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

5. D.S. Bernstein and D.C. Hyland, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation of Distributed Parameter Systems", AIAA Dynamics Specialists Conf., Palm Springs, CA, May 1984.

6. D.C. Hyland and D.S.Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation", IEEE Trans. Automat. Contr. (to appear).

7. D.C. Hyland and D.S. Bernstein, "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore", 23rd Conf. Dec. Contr., Las Vegas, NV, Dec. 1984.

8. C.S. Sims, "Reduced-Order Modelling and Filtering", in Control and Dynamic Systems, C.T. Leondes, ed., Vol. 18, pp. 55-103, 1982.

9. D.A. Wilson and R.N. Mishra, "Design of Low Order Estimators Using Reduced Models", Int. J. Control, Vol. 23, pp. 447-456, 1979.

10. K. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, New York, 1972.

11. W.M. Wonham, Linear Multivariable Control: A Geometric Approach, Springer-Verlag, New York, 1974.

12. S.L. Campbell and C.D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

13. A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse", SIAM J. Appl. Math., Vol. 17, pp. 434-440, 1969.

14. E. Kreindler and A. Jameson, "Conditions for Nonnegativeness of Partitioned Matrices", IEEE Trans. on Auto. Contr., Vol. AC-17, pp. 147-8, 1972.

15. P. J. McLane, "Optimal Linear Filtering for Linear Systems with State-Dependent Noise", Int. J. Contr., Vol. 10, pp. 42-51, 1969.

END

FILMED

1-86

DTIC